



09-924197

# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 125025**

**TO: Karen A Lacourciere**  
**Location: rem/2d15/2c18**  
**Art Unit: 1635**  
**Monday, June 21, 2004**

**Case Serial Number: 09/924197**

**From: Mary Jane Ruhl**  
**Location: Biotech-Chem Library**  
**Remsen 1-A-62**  
**Phone: 571-272-2524**

**maryjane.ruhl@uspto.gov**

### **Search Notes**

Examiner Lacourciere,

Here are the results for your recent search request.

Please feel free to contact me if you have any questions about these results.

Thank you for using STIC services. We appreciate the opportunity to serve you.

Sincerely,

Mary Jane Ruhl  
Technical Information Specialist  
STIC  
Remsen 1-A-62  
Ext. 22524

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 09:52:48 ; Search time 2214 Seconds  
(without alignments)  
12046.568 Million cell updates/sec

Title: US-09-924-197-1

Perfect score: 5822

Sequence: 1 ctggcagcaggtttcccg.....tgccgattcattaatgcag 5822

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3017426 seqs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgm2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgm2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgm2\_6/ptodata/2/pubpna/US05\_NEW\_PUB.seq.\*
- 4: /cgm2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgm2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgm2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgm2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgm2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgm2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgm2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgm2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgm2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgm2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 14: /cgm2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 15: /cgm2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 16: /cgm2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 17: /cgm2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 18: /cgm2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 19: /cgm2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5822	100.0	5822	13	US-09-924-197-1
2	2768	47.5	4773	10	US-09-991-209-32
3	2768	47.5	4950	10	US-09-991-209-34
4	2768	47.5	4965	10	US-09-991-209-37
5	2768	47.5	4974	10	US-09-991-209-35
6	2768	47.5	5001	10	US-09-991-209-40
7	2768	47.5	5034	10	US-09-991-209-33
8	2768	47.5	5164	10	US-09-991-209-36
9	2768	47.5	5277	10	US-09-991-209-25
10	2768	47.5	5295	10	US-09-991-209-38
11	2768	47.5	5327	10	US-09-991-209-27
12	2768	47.5	5337	10	US-09-991-209-19
13	2768	47.5	5337	10	US-09-991-209-23
14	2768	47.5	5337	10	US-09-991-209-31

15	2768	47.5	5338	10	US-09-991-209-15	Sequence 15, Appl
16	2768	47.5	5338	10	US-09-991-209-29	Sequence 29, Appl
17	2768	47.5	5345	10	US-09-991-209-17	Sequence 17, Appl
18	2768	47.5	5387	10	US-09-991-209-41	Sequence 41, Appl
19	2768	47.5	5395	10	US-09-991-209-21	Sequence 21, Appl
20	2683.4	46.1	3448	15	US-10-128-590-6	Sequence 6, Appl
21	2683.4	46.1	3448	15	US-10-128-587A-6	Sequence 6, Appl
22	2683.4	46.1	3448	16	US-10-128-578B-6	Sequence 6, Appl
23	2675.4	46.0	3018	13	US-09-486-142-7	Sequence 7, Appl
24	2670.2	45.9	3357	15	US-10-161-403-96	Sequence 96, Appl
25	2552	43.8	10078	14	US-10-033-190-3	Sequence 3, Appl
26	2549.4	43.8	2962	15	US-10-033-399B-10	Sequence 10, Appl
27	2549.4	43.8	3057	15	US-10-033-399B-23	Sequence 23, Appl
28	2549.4	43.8	3093	15	US-10-033-399B-7	Sequence 7, Appl
29	2549.4	43.8	5251	14	US-10-153-159-17	Sequence 17, Appl
30	2549.4	43.8	5251	15	US-10-153-176-17	Sequence 17, Appl
31	2549.4	43.8	5251	15	US-10-443-134A-17	Sequence 17, Appl
32	2549	43.8	2958	15	US-10-220-462-2	Sequence 2, Appl
33	2549	43.8	2958	15	US-10-220-262-3	Sequence 3, Appl
34	2549	43.8	2958	15	US-10-220-262-4	Sequence 4, Appl
35	2549	43.8	3351	16	US-10-014-099F-72	Sequence 72, Appl
36	2549	43.8	4754	16	US-10-014-099F-80	Sequence 80, Appl
37	2549	43.8	4773	16	US-10-014-099F-81	Sequence 81, Appl
38	2549	43.8	4831	16	US-10-014-099F-99	Sequence 99, Appl
39	2548.8	43.8	8858	15	US-10-378-393-1	Sequence 1, Appl
40	2538.2	43.6	10597	15	US-10-057-108-10	Sequence 10, Appl
41	2538.2	43.6	10599	15	US-10-057-108-11	Sequence 11, Appl
42	2538.2	43.6	12482	15	US-10-057-108-7	Sequence 7, Appl
43	2531.2	43.5	5175	15	US-10-057-108-6	Sequence 6, Appl
44	2531.2	43.5	8426	15	US-10-163-899-1	Sequence 1, Appl
45	2515	43.2	13910	10	US-09-919-901-1	Sequence 1, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-924-197-1  
; Sequence 1, Application US/09924197  
; Publication No. US20030018993A1  
; GENERAL INFORMATION:  
; APPLICANT: Gutterson, Neal  
; APPLICANT: Oeller, Paul  
; TITLE OF INVENTION: Improved Methods of Gene Silencing Using Inverted  
; TITLE OF INVENTION: Repeat Sequences  
; FILE REFERENCE: 012176-010810US  
; CURRENT APPLICATION NUMBER: US/09/924,197  
; CURRENT FILING DATE: 2001-08-07  
; PRIOR APPLICATION NUMBER: 60/225,508  
; PRIOR FILING DATE: 2000-08-15  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 5822  
; TYPE: DNA  
; ORGANISM: Agrobacterium tumefaciens  
US-09-924-197-1

Query Match	100.0%	Score 5822;	DB 13;	Length 5822;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches 5822;	Conservative	0;		Gaps 0;
Qy	1	CTGGCAGCAGAGTTTCCCGACTGGAAAGCGGCGAGTGGCGCAACGCAATTAATGTGAG	60	
Db	1	CTGGCAGCAGAGTTTCCCGACTGGAAAGCGGCGAGTGGCGCAACGCAATTAATGTGAG	60	
Qy	61	TTAGCTCACTCATTAGGCAACCCAGGCTTACATTTATGCTTCCGGTCGTTGTTG	120	
Db	61	TTAGCTCACTCATTAGGCAACCCAGGCTTACATTTATGCTTCCGGTCGTTGTTG	120	
Qy	121	TGGATTTCTGCGCGATACCAATTTTACACAGGAACAGCTATGACCATGATTATGCCCAA	180	
Db	121	TGGAATTGTGCGCGATACCAATTTTACACAGGAACAGCTATGACCATGATTATGCCCAA	180	

181 GCTATTAGGTGACACTATAGAAATACCTCAAGCTATGCAATCCAAAGCGTTGGAGCTCTCC 240  
181 GCTATTAGGTGACACTATAGAAATACCTCAAGCTATGCAATCCAAAGCGTTGGAGCTCTCC 240  
241 CATATGGTCCACCTGCAGGCGCGCACTAGTATGCTTAGATCTCTCGAGTGGAGCTTAAT 300  
241 CATATGGTCCACCTGCAGGCGCGCACTAGTATGCTTAGATCTCTCGAGTGGAGCTTAAT 300  
301 TCTCAGTCCAAAGCCTCAACAAGGTCAAGGTACAGAGTCTCCAAACCATTAGCCAAAGC 360  
301 TCTCAGTCCAAAGCCTCAACAAGGTCAAGGTACAGAGTCTCCAAACCATTAGCCAAAGC 360  
361 TACAGGAGATCAATGAAGAACTCTCAATCAAGTAACTACTGTTCCAGCACATGCATCA 420  
361 TACAGGAGATCAATGAAGAACTCTCAATCAAGTAACTACTGTTCCAGCACATGCATCA 420  
421 TGGTCAGTAAAGTTTCAGAAAAGACATCCACCGAAGACTTAAAGTTAGTGGGATCTTTG 480  
421 TGGTCAGTAAAGTTTCAGAAAAGACATCCACCGAAGACTTAAAGTTAGTGGGATCTTTG 480  
481 AAAGTAATCTTGTCAACATCGAGCAGCTGGCTTGTGGGACCAAGAAATGGT 540  
481 AAAGTAATCTTGTCAACATCGAGCAGCTGGCTTGTGGGACCAAGAAATGGT 540  
541 GCAGAAATGTTAGGCGCACCTACCAAAAGCATCTTTGGCTTTATTCGAAAGATAAGCAG 600  
541 GCAGAAATGTTAGGCGCACCTACCAAAAGCATCTTTGGCTTTATTCGAAAGATAAGCAG 600  
601 ATTCTCTAGTACAAAGTGGGGAACAAATAAAGTGGAAAGAGAGCTGTCTGACAGCCAC 660  
601 ATTCTCTAGTACAAAGTGGGGAACAAATAAAGTGGAAAGAGAGCTGTCTGACAGCCAC 660  
661 TCACTAATGCGTATGACAAAGCAGTACGACCAACAAAGAAATTTAGCTCAGGAT 720  
661 TCACTAATGCGTATGACAAAGCAGTACGACCAACAAAGAAATTTAGCTCAGGAT 720  
721 TTAGCAGCATTCAGATTTGGTTTCAATCAACAGGTAGAGCCATATCACTTTATTCAAA 780  
721 TTAGCAGCATTCAGATTTGGTTTCAATCAACAGGTAGAGCCATATCACTTTATTCAAA 780  
781 TTGATATGCGCAAAACCAAGAGGAACTCCCATCTCCAAAGGTTTGTAAAGGAATTC 840  
781 TTGATATGCGCAAAACCAAGAGGAACTCCCATCTCCAAAGGTTTGTAAAGGAATTC 840  
841 CAGTCCAAAGCCTCAACAAGGTCAAGGTACAGAGTCTCCAAACCATTAGCCAAAGCTAC 900  
841 CAGTCCAAAGCCTCAACAAGGTCAAGGTACAGAGTCTCCAAACCATTAGCCAAAGCTAC 900  
901 AGGAGATCAATGAAGAACTCTCAATCAAGTAACTACTGTTCCAGCACATGCATCATGG 960  
901 AGGAGATCAATGAAGAACTCTCAATCAAGTAACTACTGTTCCAGCACATGCATCATGG 960  
961 TCAGTAAAGTTTCAGAAAAGACATCCACCGAAGACTTAAAGTTAGTGGGATCTTTGAAA 1020  
961 TCAGTAAAGTTTCAGAAAAGACATCCACCGAAGACTTAAAGTTAGTGGGATCTTTGAAA 1020  
1021 GTAACTTTGTCAACATCGAGCAGCTGGCTTGTGGGACCAAGAAATGGTGA 1080  
1021 GTAACTTTGTCAACATCGAGCAGCTGGCTTGTGGGACCAAGAAATGGTGA 1080  
1081 GAATTTGTAGGCGCACCTACCAAAAGCATCTTTGGCTTTATTCGAAAGATAAGCAGATT 1140  
1081 GAATTTGTAGGCGCACCTACCAAAAGCATCTTTGGCTTTATTCGAAAGATAAGCAGATT 1140  
1141 CCTCTAGTACAAAGTGGGGAACAAATAAAGTGGAAAGAGCTGTCTGACAGCCCACTCA 1200  
1141 CCTCTAGTACAAAGTGGGGAACAAATAAAGTGGAAAGAGCTGTCTGACAGCCCACTCA 1200  
1201 CTAATGCGTATGACGAACGCGATGACGACCAAAAGAAATTCCTCTATATAGAAGGCA 1260  
1201 CTAATGCGTATGACGAACGCGATGACGACCAAAAGAAATTCCTCTATATAGAAGGCA 1260

1261 TTCAATTTCCATTTGAAGGACACAGAAATTTGCTACATTTGTTTCAAACTTCAAAATAT 1320  
1261 TTCAATTTCCATTTGAAGGACACAGAAATTTGCTACATTTGTTTCAAACTTCAAAATAT 1320  
1321 TATTCAATTTATTGTCAGCTTTCAAACTCTTTGTTTCTGTTGTTGTTGTTGTTGTTGTT 1380  
1321 TATTCAATTTATTGTCAGCTTTCAAACTCTTTGTTTCTGTTGTTGTTGTTGTTGTTGTT 1380  
1381 AAAACCATGTTCTTAAACCAAGAAATTTATCTTCTCAAGCAATCACCTTTTCAGGTCCA 1440  
1381 AAAACCATGTTCTTAAACCAAGAAATTTATCTTCTCAAGCAATCACCTTTTCAGGTCCA 1440  
1441 TGCAGATCTTCTATTTTCAGTAAAGATTTTGGATCTTGAAGCATCTAGTAAATTTTCA 1500  
1441 TGCAGATCTTCTATTTTCAGTAAAGATTTTGGATCTTGAAGCATCTAGTAAATTTTCA 1500  
1501 GACTCAAAAGATAGAGGCTTTGGATTTGCTTTTGTATAGTCTTCAAAATTTAGTTTGGGA 1560  
1501 GACTCAAAAGATAGAGGCTTTGGATTTGCTTTTGTATAGTCTTCAAAATTTAGTTTGGGA 1560  
1561 GGAGGAGAACTATCAATGGCAATGGCAAGATATGGTGGCCAAAGTTCTTGCAGAAATTAAT 1620  
1561 GGAGGAGAACTATCAATGGCAATGGCAAGATATGGTGGCCAAAGTTCTTGCAGAAATTAAT 1620  
1621 AAATCACTGTCATGCGAGGATGCAACAAAGGCTTAAACCTTCTGGAATTTGCAAAATTTG 1680  
1621 AAATCACTGTCATGCGAGGATGCAACAAAGGCTTAAACCTTCTGGAATTTGCAAAATTTG 1680  
1681 AAAGTGAATTAATCTAAAGAGTAAATAATGCAACAAATTCATATCAAAATTTGAGTCATGC 1740  
1681 AAAGTGAATTAATCTAAAGAGTAAATAATGCAACAAATTCATATCAAAATTTGAGTCATGC 1740  
1741 ACTAATGTTGTAGCTTTCAAAATTTGATGATCAATGCTTTCAGCAAGAGCCCAAAATCTGAT 1800  
1741 ACTAATGTTGTAGCTTTCAAAATTTGATGATCAATGCTTTCAGCAAGAGCCCAAAATCTGAT 1800  
1801 GAGTCCATGATCAAAATCTCAATATATCAATATCTGATATCTGATATCTGATATCTGAT 1860  
1801 GAGTCCATGATCAAAATCTCAATATATCTGATATCTGATATCTGATATCTGATATCTGAT 1860  
1861 GATGATGTTTCAATTTGTTCTGGATCTCAAAATTTGTCAGGCGCAAAATTTACTTGT 1920  
1861 GATGATGTTTCAATTTGTTCTGGATCTCAAAATTTGTCAGGCGCAAAATTTACTTGT 1920  
1921 GGTCCAGGTCACTGATTAAGTATTTGGAAGCTTAGGATCTGGAATTTGAGAGCTTATGTG 1980  
1921 GGTCCAGGTCACTGATTAAGTATTTGGAAGCTTAGGATCTGGAATTTGAGAGCTTATGTG 1980  
1981 TCTAATGTTTACTGTAATGAAGCCAAATTTATCGGTGCCGAAATTTGAGTCAAG 2040  
1981 TCTAATGTTTACTGTAATGAAGCCAAATTTATCGGTGCCGAAATTTGAGTCAAG 2040  
2041 ACTTGCGAGGAGGATCTGGAAGCTAGCAACATCAAAATTTCTGAAATTTGGAATTCGAA 2100  
2041 ACTTGCGAGGAGGATCTGGAAGCTAGCAACATCAAAATTTCTGAAATTTGGAATTCGAA 2100  
2101 GAGTTTAACTATCCCATTAATTTAGACCAAACTTTATTTGATCGAGTTGAAACCATGTATA 2160  
2101 GAGTTTAACTATCCCATTAATTTAGACCAAACTTTATTTGATCGAGTTGAAACCATGTATA 2160  
2161 CAAACAGTTTTTCAGCAGTTCAAGTGAATAATTTGGTGTATGAGAAATATCAAGGGCACAAGT 2220  
2161 CAAACAGTTTTTCAGCAGTTCAAGTGAATAATTTGGTGTATGAGAAATATCAAGGGCACAAGT 2220  
2221 GCAACAAAGTGGCCATAAATTTGATTCAGCACCAAACTTCCATGTTGAGGAAATTTATA 2280  
2221 GCAACAAAGTGGCCATAAATTTGATTCAGCACCAAACTTCCATGTTGAGGAAATTTATA 2280  
2281 ATGAGAAATATAAATTTAGTGGGGAAGTGGAAACCAATCAGAGGCTAGCTGCAAAAT 2340  
2281 ATGAGAAATATAAATTTAGTGGGGAAGTGGAAACCAATCAGAGGCTAGCTGCAAAAT 2340  
2341 GTCCATTTTAAACAAATGCTGAACATGTTTACACCACACTGCACTTCACTAGAAATTTTCAGAG 2400

Db 2341 GTCCATTTTAAACATGCTGAACATGTTTACACACACTGCACTTCTCCTAGAAAATTTTCAGAG 2400  
Qy 2401 GATCAAGCTCTTTTGTATTAATTAATGTCAGAGTGCATCTAGTAAACATAGATGACACC 2460  
Db 2401 GATCAAGCTCTTTTGTATTAATTAATGTCAGAGTGCATCTAGTAAACATAGATGACACC 2460  
Qy 2461 GCGCGGATTAATTTATCTAGTTTGGCGGCTATATTTTGTCTTATCGCGTATTAATATG 2520  
Db 2461 GCGCGGATTAATTTATCTAGTTTGGCGGCTATATTTTGTCTTATCGCGTATTAATATG 2520  
Qy 2521 ATAATTCGGGACTCTAATCAGAAAAACCCATCTCATATAATTAACGTCATGCAATTAATG 2580  
Db 2521 ATAATTCGGGACTCTAATCAGAAAAACCCATCTCATATAATTAACGTCATGCAATTAATG 2580  
Qy 2581 TAATTAATTAACATGCTTAACGTAATTTCAACAGAAATTAATATGATTAATCATCGCAAGACGG 2640  
Db 2581 TAATTAATTAACATGCTTAACGTAATTTCAACAGAAATTAATATGATTAATCATCGCAAGACGG 2640  
Qy 2641 CAACAGGATTCATCTTAAGAAACCTTTATTCGCCAAATGTTTGAACATCTGCTTGACTTA 2700  
Db 2641 CAACAGGATTCATCTTAAGAAACCTTTATTCGCCAAATGTTTGAACATCTGCTTGACTTA 2700  
Qy 2701 GAGGATCTTCAATTTTACTGTGCAACATTTCTGCTGAATTTGTTTATCATAATAA 2760  
Db 2701 GAGGATCTTCAATTTTACTGTGCAACATTTCTGCTGAATTTGTTTATCATAATAA 2760  
Qy 2761 TTGGTTTCGTTAAATTTGTAATTTGCTTCTTATTTTGAACCAATCAAAAGCGGCTACGG 2820  
Db 2761 TTGGTTTCGTTAAATTTGTAATTTGCTTCTTATTTTGAACCAATCAAAAGCGGCTACGG 2820  
Qy 2821 ATCTTCCTAGAGTCAGCAGATCGTTCAACATTTGGCAATAAAGTTTCTTAAGATTGAA 2880  
Db 2821 ATCTTCCTAGAGTCAGCAGATCGTTCAACATTTGGCAATAAAGTTTCTTAAGATTGAA 2880  
Qy 2881 TCCGTTGCGGCTCTCGATGATTTATCATATAATTTCTGTTGAATTTACGTTAAGCATGT 2940  
Db 2881 TCCGTTGCGGCTCTCGATGATTTATCATATAATTTCTGTTGAATTTACGTTAAGCATGT 2940  
Qy 2941 AATAAATTAACATGTAATGATGATGATTTATTTATGAGATGGGTTTTTATGATTTAGAGTCCC 3000  
Db 2941 AATAAATTAACATGTAATGATGATGATTTATTTATGAGATGGGTTTTTATGATTTAGAGTCCC 3000  
Qy 3001 GCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3060  
Db 3001 GCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3060  
Qy 3061 ATCCGCGCGGCTCATCTAATTTACTAGATCGACCTGCGAGGATGGATCCGCGCGCG 3120  
Db 3061 ATCCGCGCGGCTCATCTAATTTACTAGATCGACCTGCGAGGATGGATCCGCGCGCG 3120  
Qy 3121 ATCCGAGCTCGGCGCCAAATTCGCCCTATAGTGTGATTTATTAATTTCACTGGCGGCTCGT 3180  
Db 3121 ATCCGAGCTCGGCGCCAAATTCGCCCTATAGTGTGATTTATTAATTTCACTGGCGGCTCGT 3180  
Qy 3181 TTTTAAACGCTGTGATCGGGAAACCCCTGGGTTTACCCAACTTAATCGCTTTCAGGACACA 3240  
Db 3181 TTTTAAACGCTGTGATCGGGAAACCCCTGGGTTTACCCAACTTAATCGCTTTCAGGACACA 3240  
Qy 3241 TCCCCCTTTGCGAGCTGGGCTAATAGGAGAGGCGCGACCGATCGCCCTTCCCAACA 3300  
Db 3241 TCCCCCTTTGCGAGCTGGGCTAATAGGAGAGGCGCGACCGATCGCCCTTCCCAACA 3300  
Qy 3301 GTTCGCGAGCTGAAATGGCGAATGGACGCGCCCTGTAGCGCGCATTAAGCGCGCGGGT 3360  
Db 3301 GTTCGCGAGCTGAAATGGCGAATGGACGCGCCCTGTAGCGCGCATTAAGCGCGCGGGT 3360  
Qy 3361 GTGTGTGTTACGCGAGCGGTGACCGCTACACTTCCGAGCGGCTTAGCGCGGCTCTCTTC 3420  
Db 3361 GTGTGTGTTACGCGAGCGGTGACCGCTACACTTCCGAGCGGCTTAGCGCGGCTCTCTTC 3420  
Qy 3421 GCTTTCTTCCCTTCTCTGCGCAAGTTCGCGGCTTCCCGCTCAAGCTCTAAATCGG 3480

Db 3421 GCTTTCTTCCCTTCTCTGCGCAAGTTCGCGGCTTCCCGCTCAAGCTCTAAATCGG 3480  
Qy 3481 GGGCTCCCTTTAGGTTTCGATTTAGAGCTTTTACGCACTCGACCGCAAAAAAATCTTGAT 3540  
Db 3481 GGGCTCCCTTTAGGTTTCGATTTAGAGCTTTTACGCACTCGACCGCAAAAAAATCTTGAT 3540  
Qy 3541 TTGGGTCGATGTTTCAAGCTAGTGGGCCATCGCCTGATAGACGGTTTTCGCGCTTTGAGC 3600  
Db 3541 TTGGGTCGATGTTTCAAGCTAGTGGGCCATCGCCTGATAGACGGTTTTCGCGCTTTGAGC 3600  
Qy 3601 TTGAGTTCAGCTCTTTTAAATAGTGGACTCTTGTTCAAAACTGGAAACAACTCAACCT 3660  
Db 3601 TTGAGTTCAGCTCTTTTAAATAGTGGACTCTTGTTCAAAACTGGAAACAACTCAACCT 3660  
Qy 3661 ATCTCGGCTATTTCTTTTGAATTAAGGAAATTTTGGCGATTTTCGCGCTATTTGGTTAAAA 3720  
Db 3661 ATCTCGGCTATTTCTTTTGAATTAAGGAAATTTTGGCGATTTTCGCGCTATTTGGTTAAAA 3720  
Qy 3721 AATGAGCTGATTTTAAACAAATTAATTAACGCAATTTTAAACAAATTAATTAACGTTTACAAT 3780  
Db 3721 AATGAGCTGATTTTAAACAAATTAATTAACGCAATTTTAAACAAATTAATTAACGTTTACAAT 3780  
Qy 3781 TCGCTCATGCGGTATTTCTTTCAGCATCTGTCGGGTATTTTCAACCGCATACAGT 3840  
Db 3781 TCGCTCATGCGGTATTTCTTTCAGCATCTGTCGGGTATTTTCAACCGCATACAGT 3840  
Qy 3841 GGCATTTTTCGCGGAAATGTCGCGAAACCCCTATTTTGTATTTTCTTAAATACATTTCA 3900  
Db 3841 GGCATTTTTCGCGGAAATGTCGCGAAACCCCTATTTTGTATTTTCTTAAATACATTTCA 3900  
Qy 3901 AATATGATCGCTCATGAGACATTAACCTGATAAATGCTTCAATAATTTTGAAGAGG 3960  
Db 3901 AATATGATCGCTCATGAGACATTAACCTGATAAATGCTTCAATAATTTTGAAGAGG 3960  
Qy 3961 AAGAGTATGAGTATTTCAACATTTCCGCTGTCGCCCTTATTCCTTTTTCGCGCATTTTGC 4020  
Db 3961 AAGAGTATGAGTATTTCAACATTTCCGCTGTCGCCCTTATTCCTTTTTCGCGCATTTTGC 4020  
Qy 4021 CTTCTCTTTTTCGTCACCCAGAAAACGCTGCTGTAAGAGTAAAGATGCTGAAGATCAGTTG 4080  
Db 4021 CTTCTCTTTTTCGTCACCCAGAAAACGCTGCTGTAAGAGTAAAGATGCTGAAGATCAGTTG 4080  
Qy 4081 GGTGACAGGTGGTATCATCGAATCTCAACAGCGGTAAAGATGCTTGAAGTTT 4140  
Db 4081 GGTGACAGGTGGTATCATCGAATCTCAACAGCGGTAAAGATGCTTGAAGTTT 4140  
Qy 4141 CGCCCCGAGAAACGTTTCCAAATGATGAGCACTTTTAAAGTTTCTGCTATGTTGGCGGTA 4200  
Db 4141 CGCCCCGAGAAACGTTTCCAAATGATGAGCACTTTTAAAGTTTCTGCTATGTTGGCGGTA 4200  
Qy 4201 TTATCCCGTATTTGACCGCGGCAAGAGCACTCGGTCGCCCATACACTATTTCTCAGAT 4260  
Db 4201 TTATCCCGTATTTGACCGCGGCAAGAGCACTCGGTCGCCCATACACTATTTCTCAGAT 4260  
Qy 4261 GACTTGTGAGTACTCACCAGTCAAGAAAGCATCTTACGATGGATGGATGAGTAAGA 4320  
Db 4261 GACTTGTGAGTACTCACCAGTCAAGAAAGCATCTTACGATGGATGGATGAGTAAGA 4320  
Qy 4321 GAATTAATGAGTGTCTGCTCAATAACCATGAGTGAATAACACTGCGGCCAACTTACTTCTGA 4380  
Db 4321 GAATTAATGAGTGTCTGCTCAATAACCATGAGTGAATAACACTGCGGCCAACTTACTTCTGA 4380  
Qy 4381 ACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTCGACAACTATGCGGGATCATGTAAT 4440  
Db 4381 ACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTCGACAACTATGCGGGATCATGTAAT 4440  
Qy 4441 CGCCTTGCATGTTGGGAAACCGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACC 4500  
Db 4441 CGCCTTGCATGTTGGGAAACCGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACC 4500  
Qy 4501 ACGATGCTCTGATGCAATGGCAACAAACGTTGCGCAAACTATTTAACTGCGGAACTACTTACT 4560  
Db 4501 ACGATGCTCTGATGCAATGGCAACAAACGTTGCGCAAACTATTTAACTGCGGAACTACTTACT 4560



```
QY 4561 CTAGCTTCCGGCAACAAATTAATAGACTGGATGGAGCGGATAAAGTTGACAGACCACTT 4620
DB 4561 CTAGCTTCCGGCAACAAATTAATAGACTGGATGGAGCGGATAAAGTTGACAGACCACTT 4620
QY 4621 CTGGCTCGGCCCTTCCGGCTGGCTGGTTTATTTGCTGATAAATCTTGAGCGCGGTGAGCGT 4680
DB 4621 CTGGCTCGGCCCTTCCGGCTGGCTGGTTTATTTGCTGATAAATCTTGAGCGCGGTGAGCGT 4680
QY 4681 GGCTCTCGCGGTATCATATGACGACATGGGGCCAGATGGTAAAGCCCTCCCGTATCGTAGTT 4740
DB 4681 GGCTCTCGCGGTATCATATGACGACATGGGGCCAGATGGTAAAGCCCTCCCGTATCGTAGTT 4740
QY 4741 ATCTACAGACGGGGAGTCAGGCAACTATGCGATGACGAATAATAGACAGATCGCTGAGATA 4800
DB 4741 ATCTACAGACGGGGAGTCAGGCAACTATGCGATGACGAATAATAGACAGATCGCTGAGATA 4800
QY 4801 GGTGCTCTACTGATTAAGCAATTTGTAAGTCTGACAGCAAGTTTACTCATATATACTTTAG 4860
DB 4801 GGTGCTCTACTGATTAAGCAATTTGTAAGTCTGACAGCAAGTTTACTCATATATACTTTAG 4860
QY 4861 ATTGATTTAAATCTTCAATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAAT 4920
DB 4861 ATTGATTTAAATCTTCAATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAAT 4920
QY 4921 CTCATGACCAAAATCCCTTAACGTAGTTTCTGTTCCACTGACGCTCAGACCCCGTAGAA 4980
DB 4921 CTCATGACCAAAATCCCTTAACGTAGTTTCTGTTCCACTGACGCTCAGACCCCGTAGAA 4980
QY 4981 AAGATCAAGGATCTCTGAGATCCTTTTCTGCGGCTAATCTGCTGCTTGCAAAACA 5040
DB 4981 AAGATCAAGGATCTCTGAGATCCTTTTCTGCGGCTAATCTGCTGCTTGCAAAACA 5040
QY 5041 AAAAAACACCGCTACAGCGGTGGTTGTTGTCGGGATCAAGAGCTACCAACTCTTTT 5100
DB 5041 AAAAAACACCGCTACAGCGGTGGTTGTTGTCGGGATCAAGAGCTACCAACTCTTTT 5100
QY 5101 CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAATCTGCTCTTAGTGTAGCGG 5160
DB 5101 CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAATCTGCTCTTAGTGTAGCGG 5160
QY 5161 TAGTTAGGCAACCACTTCAAGAACTCTGTAGCAACCGCTACATCTGCTCTGCTAAATC 5220
DB 5161 TAGTTAGGCAACCACTTCAAGAACTCTGTAGCAACCGCTACATCTGCTCTGCTAAATC 5220
QY 5221 CTGTTACCAAGTGGCTGCTCAGTGGCGAFAAGTGTGTTCTTACCAGGTTGGACTCAAGA 5280
DB 5221 CTGTTACCAAGTGGCTGCTCAGTGGCGAFAAGTGTGTTCTTACCAGGTTGGACTCAAGA 5280
QY 5281 CGATAGTTACCGGATAAGCGGAGCGGTGGGCTGAAACGGGGGGTTGCTGCACACAGCCC 5340
DB 5281 CGATAGTTACCGGATAAGCGGAGCGGTGGGCTGAAACGGGGGGTTGCTGCACACAGCCC 5340
QY 5341 AGCTTGAGGCGAAGCACTACACCGAACTGAGATACCTACAGCGCTGAGCTATGAGAAGC 5400
DB 5341 AGCTTGAGGCGAAGCACTACACCGAACTGAGATACCTACAGCGCTGAGCTATGAGAAGC 5400
QY 5401 GCCACGCTTCCGAAAGGAGAAAGGCGGACAGGTATCOGTTAAGCGGCGAGGGTCGGAACA 5460
DB 5401 GCCACGCTTCCGAAAGGAGAAAGGCGGACAGGTATCOGTTAAGCGGCGAGGGTCGGAACA 5460
QY 5461 GGNAGGCGCAAGAGGAGCTTCCAGGGGGAACGCTCTGATCTTTATAGTCTGTCGGG 5520
DB 5461 GGNAGGCGCAAGAGGAGCTTCCAGGGGGAACGCTCTGATCTTTATAGTCTGTCGGG 5520
QY 5521 TTTTCGCCACTCTGACTGAGCGTCAATTTTGTGATGCTGCTCAGGGGGCGGAGCCTA 5580
DB 5521 TTTTCGCCACTCTGACTGAGCGTCAATTTTGTGATGCTGCTCAGGGGGCGGAGCCTA 5580
QY 5581 TGGAAAAACGCCAGCAACCGCGCTTTTATACGTTCTGCGCTTTTGTGCGCTTTTGTCT 5640
DB 5581 TGGAAAAACGCCAGCAACCGCGCTTTTATACGTTCTGCGCTTTTGTGCGCTTTTGTCT 5640
```

```
QY 5641 CACATCTTCTTCTCCTGCTATCCCTGATCTCTGTGATAAACCGTATTACCGCCTTTGAG 5700
DB 5641 CACATCTTCTTCTCCTGCTATCCCTGATCTCTGTGATAAACCGTATTACCGCCTTTGAG 5700
QY 5701 TGAAGCTGATACCGCTCGCGCAGCGAAACGACCGAGCGCAGCGAGTCAGTGAGCGAGAA 5760
DB 5701 TGAAGCTGATACCGCTCGCGCAGCGAAACGACCGAGCGCAGCGAGTCAGTGAGCGAGAA 5760
QY 5761 GCGGAAGAGCGCCCAATAGCGAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGC 5820
DB 5761 GCGGAAGAGCGCCCAATAGCGAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGC 5820
QY 5821 AG 5822
DB 5821 AG 5822

RESULT 2
US-09-991-209-32
; Sequence 32, Application US/09991209
; Publication No. US20030024009A1
; GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Langdon, Timothy
; APPLICANT: Morse, Phillip
; TITLE OF INVENTION: Manipulation of the Phenolic Acid
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted
; FILE REFERENCE: GC648-2
; CURRENT APPLICATION NUMBER: US/09/991,209
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/249,608
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 4773
; TPB: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pG6 vector
US-09-991-209-32

Query Match 47.5%; Score 2768; DB 10; Length 4773;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 2895; Conservative 0; Mismatches 25; Indels 72; Gaps 4;

QY 2835 AAGCAGATCGTTCAAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCTCTGTGCGGTC 2894
DB 778 AAGCAGATCGTTCAAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCTCTGTGCGGTC 837
QY 2895 TTGCGATGATTATCATATAATTTCTGTTGAATTAAGTTAAGCATGTAATAATTAACATGT 2954
DB 838 TTGCGATGATTATCATATAATTTCTGTTGAATTAAGTTAAGCATGTAATAATTAACATGT 897
QY 2955 AATGCAATGCTTTTATGAGATGGTTTATGATAGTCCCGCAATTAACATTT 3014
DB 898 AATGCAATGCTTTTATGAGATGGTTTATGATAGTCCCGCAATTAACATTT 957
QY 3015 AATACCGCATAGAAAACAAAATATAGCGGCAAACTAGAGATAAATATACGCGCGGTGT 3074
DB 958 AATACCGCATAGAAAACAAAATATAGCGGCAAACTAGAGATAAATATACGCGCGGTGT 1017
QY 3075 CATCTATGTTACTAGATTCGACCTGAGGCAATGGGATCGCGGCGCATGCGACGTGGGC 3134
DB 1018 CATCTATGTTACTAGATTCGATAGCTTCTAGAGCGCGGTGGAGC-----T 1064
QY 3135 CCAATTCGCGCTATAGTGTGATTAAC---AATTCACCTGCGCGCTGTTTACACGTC 3191
DB 1065 CCAATTCGCGCTATAGTGTGATTAAC---AATTCACCTGCGCGCTGTTTACACGTC 1124
QY 3192 GTGACTGGGAAACCCCTGGCGTTTACCAACTTAATGCGCTTGACGACATATCCCTTCG 3251
```

Db 1125 GTGACTGGGAAAAACCTCGCGTTACCCCACTTAATCGCCTTGCAGCACATCCCCCTTTCG 1184  
Qy 3252 CCAGCTGCGGTAATAGCGAAGAGGCCCGACCGATCGCCCTTCCAAACAGTTGGCGAGCC 3311  
Db 1185 CCAGCTGCGGTAATAGCGAAGAGGCCCGACCGATCGCCCTTCCAAACAGTTGGCGAGCC 1244  
Qy 3312 TGAATGGCGAAT-GGACGCGCCTCTAGCGCGCATTAAGCGCGCGGCTGTGGTGGTTA 3370  
Db 1245 TGAATGGCGAATGGAGCGCGCCTGTAGCGCGCATTAAGCGCGCGGCTGTGGTGGTTA 1304  
Qy 3371 CGCGCAGCGTGAACCGCTACACTTGCACGCGCCTAGCGCCCGCTCTCTTTTCGCTTCTTCC 3430  
Db 1305 CGCGCAGCGTGAACCGCTACACTTGCACGCGCCTAGCGCCCGCTCTCTTTTCGCTTCTTCC 1364  
Qy 3431 CTTCTCTTCTCGCACGTTTCGCGGCTTTCGCGCTTCAAGCTCTAAATCGGGGCTCCCTT 3490  
Db 1365 CTTCTCTTCTCGCACGTTTCGCGGCTTTCGCGCTTCAAGCTCTAAATCGGGGCTCCCTT 1424  
Qy 3491 TAGGTTCCGATTTAGAGCTTTAGCGCACCTCGACCGCAAAAACCTGATTTGGGTGATG 3550  
Db 1425 TAGGTTCCGATTTAGAGCTTTAGCGCACCTCGACCGCAAAAACCTTGNATTAGGGTGA 1484  
Qy 3551 GTTCACTAGTGGGCCATCGCCCTGATAGAGCGTTTTCGCGCTTTCGACGTTGGAGTCCA 3610  
Db 1485 GTTCACTAGTGGGCCATCGCCCTGATAGAGCGTTTTCGCGCTTTCGACGTTGGAGTCCA 1544  
Qy 3611 CGTTCTTTAATAGTGGACTCTGTTCCTCAAACTGGAAACAACTCAACCTATCTCGGTCT 3670  
Db 1545 CGTTCTTTAATAGTGGACTCTGTTCCTCAAACTGGAAACAACTCAACCTATCTCGGTCT 1604  
Qy 3671 ATTCTTTTGAATTAAGGATTTTGGCGATTTTCGCGCTATGGTTAAATAAGACTGA 3730  
Db 1605 ATTCTTTTGAATTAAGGATTTTGGCGATTTTCGCGCTATGGTTAAATAAGACTGA 1664  
Qy 3731 TTTTAAACAAATTTTAAACGCGAATTTTAAACAAATTTTAAACGCGTTCGCTGATG 3790  
Db 1665 TTTTAAACAAATTTTAAACGCGAATTTTAAACAAATTTTAAACGCGTTCGCTGATG 1715  
Qy 3791 CGGTATTTTCTCTACGCATCTGTGGGTATTTTCAACCGCATACAGTGGCGACTTTTC 3850  
Db 1716 -----AGGTGGCGACTTTTC 1729  
Qy 3851 GGGGAAATGTGCGGGAACCCCTATTTGTTTATTTTCTAAATACATTTCAATATGTATC 3910  
Db 1730 GGGGAAATGTGCGGGAACCCCTATTTGTTTATTTTCTAAATACATTTCAATATGTATC 1789  
Qy 3911 CGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAGGAAGATGA 3970  
Db 1790 CGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAGGAAGATGA 1849  
Qy 3971 GTATTCAACATTTCCGTGTGCCCTTATTCCTTTTTCGCGCATTTTTCCTTCTGTTT 4030  
Db 1850 GTATTCAACATTTCCGTGTGCCCTTATTCCTTTTTCGCGCATTTTTCCTTCTGTTT 1909  
Qy 4031 TTGCTCACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGTGCACGAG 4090  
Db 1910 TTGCTCACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGTGCACGAG 1969  
Qy 4091 TGGGTTACATCGAATCGATCTCAACAGCGTAAAGATCTTTCGAGAGTTTTCGCGCCCGAAG 4150  
Db 1970 TGGGTTACATCGAATCGATCTCAACAGCGTAAAGATCTTTCGAGAGTTTTCGCGCCCGAAG 2029  
Qy 4151 AACGTTTTCGAATGAGCACTTTTAAAGTTTCGCTATGTGGCGCGGTATTTATCCCGTA 4210  
Db 2030 AACGTTTTCGAATGAGCACTTTTAAAGTTTCGCTATGTGGCGCGGTATTTATCCCGTA 2089  
Qy 4211 TTGAGCGCGGCAAGAGCACTTCGCTCGCGCATACACTATTTCTCAGATGACTTGGTTG 4270  
Db 2090 TTGAGCGCGGCAAGAGCACTTCGCTCGCGCATACACTATTTCTCAGATGACTTGGTTG 2149  
Qy 4271 AGTACTCACAGTCAAGAAAGCATCTTAAACGATGGCATGACAGTAAGAGAAATTTATGCA 4330  
Db 2150 AGTACTCACAGTCAAGAAAGCATCTTAAACGATGGCATGACAGTAAGAGAAATTTATGCA 2209

Qy 4331 GTGCTGCCATTAACCATGAGTGATTAACACTGCGGCCAACTTACTTCTGACAAAGATCGGAG 4390  
Db 2210 GTGCTGCCATTAACCATGAGTGATTAACACTGCGGCCAACTTACTTCTGACAAAGATCGGAG 2269  
Qy 4391 GACCGAAGGAGCTAACCCCTTTTTCGCAACAATGGGGGATCATGTAACTCGCCTTGTATC 4450  
Db 2270 GACCGAAGGAGCTAACCCCTTTTTCGCAACAATGGGGGATCATGTAACTCGCCTTGTATC 2329  
Qy 4451 GTTGGGAAACCGGAGCTGAATGAAGCATATCAAAACGACGAGGTGACACAGATGCTTG 4510  
Db 2330 GTTGGGAAACCGGAGCTGAATGAAGCATATCAAAACGACGAGGTGACACAGATGCTTG 2389  
Qy 4511 TAGCAATGGCAACAACGTTGCGCAAACTATTAATCTGGCGAACTACTTACTCTAGCTTCCC 4570  
Db 2390 TAGCAATGGCAACAACGTTGCGCAAACTATTAATCTGGCGAACTACTTACTCTAGCTTCCC 2449  
Qy 4571 GGCACAAATTAATAGACTGGATGGAGCGGATAAAGTTTCAGGACCACTTCTGCGCTCGG 4630  
Db 2450 GGCACAAATTAATAGACTGGATGGAGCGGATAAAGTTTCAGGACCACTTCTGCGCTCGG 2509  
Qy 4631 CCTTTCGCGCTGGCTGTTTATTCGTATAAATCTGGAGCGGTGAGCGTGGTCTCGCG 4690  
Db 2510 CCTTTCGCGCTGGCTGTTTATTCGTATAAATCTGGAGCGGTGAGCGTGGTCTCGCG 2569  
Qy 4691 GTATCATTCGACGACTGGGCGCAGATGCTAAGCCCTCCCGTATCGTAGTTATCTACACA 4750  
Db 2570 GTATCATTCGACGACTGGGCGCAGATGCTAAGCCCTCCCGTATCGTAGTTATCTACACA 2629  
Qy 4751 CGGCGAGTCAGGCAACTATGTATGAACGAAATAGACAGATCGCTGAGATAGTGTCTCAC 4810  
Db 2630 CGGCGAGTCAGGCAACTATGTATGAACGAAATAGACAGATCGCTGAGATAGTGTCTCAC 2689  
Qy 4811 TGATTAAGCATTTGTATACTGTTCAGACCAAGTTTACTATATATATCTTTAGATTTGATTA 4870  
Db 2690 TGATTAAGCATTTGTATACTGTTCAGACCAAGTTTACTATATATATCTTTAGATTTGATTA 2749  
Qy 4871 AACTTCATTTTAAATTAAGGATCTAGGTGAAGATCCTTTTGTATATCTCATGACCA 4930  
Db 2750 AACTTCATTTTAAATTAAGGATCTAGGTGAAGATCCTTTTGTATATCTCATGACCA 2809  
Qy 4931 AACTTCATTTTAAATTAAGGATCTAGGTGAAGATCCTTTTGTATATCTCATGACCA 4990  
Db 2810 AACTTCATTTTAAATTAAGGATCTAGGTGAAGATCCTTTTGTATATCTCATGACCA 2869  
Qy 4991 GATCTTCTGAGATCTTTTTCGCGGTAATCTGCTGTGCAAAACAAACCAAC 5050  
Db 2870 GATCTTCTGAGATCTTTTTCGCGGTAATCTGCTGTGCAAAACAAACCAAC 2929  
Qy 5051 CGCTACAGCGGTGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTTCGAGAGTAA 5110  
Db 2930 CGCTACAGCGGTGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTTCGAGAGTAA 2989  
Qy 5111 CTGGCTTCAGCAGAGCGCAGATCAAAATCTGCTTCTAGTGTAGCGGTAGTGGCC 5170  
Db 2990 CTGGCTTCAGCAGAGCGCAGATCAAAATCTGCTTCTAGTGTAGCGGTAGTGGCC 3049  
Qy 5171 ACCATTTCAAGAACTCTGTAGCAGCCGCTACATACCTCGCTCTGCTATCTCTTTAC 5230  
Db 3050 ACCATTTCAAGAACTCTGTAGCAGCCGCTACATACCTCGCTCTGCTATCTCTTTAC 3109  
Qy 5231 TGGCTGCTGCGAGTGGCGATAGTGTCTTACCGGTTGGACTCAAGAGATAGTTTAC 5290  
Db 3110 TGGCTGCTGCGAGTGGCGATAGTGTCTTACCGGTTGGACTCAAGAGATAGTTTAC 3169  
Qy 5291 CGGATAAGCGCAGCGGTGCGGCTGAAACGCGGGGTTCGTGCAACAGCCAGCTTGGAGC 5350  
Db 3170 CGGATAAGCGCAGCGGTGCGGCTGAAACGCGGGGTTCGTGCAACAGCCAGCTTGGAGC 3229  
Qy 5351 GAACGACCTTACACCGAACTGAGATACCTTACAGCGTGAAGTATGAGAAAGCGCCAGCTTC 5410  
Db 3230 GAACGACCTTACACCGAACTGAGATACCTTACAGCGTGAAGTATGAGAAAGCGCCAGCTTC 3289

```
QY 5411 CCGAAGGAGAAAGCGGACAGGTATCCGTAAAGCGCAGGGTCGGACAGGAGAGCGCA 5470
DB 3290 CCGAAGGAGAAAGCGGACAGGTATCCGTAAAGCGCAGGGTCGGACAGGAGAGCGCA 3349
QY 5471 CGAGGGAGCTTCCAGGGGGAAGCGCCCTGTATCTTTATAGTCCCTGCGGTTCGCCACCC 5530
DB 3350 CGAGGGAGCTTCCAGGGGGAAGCGCCCTGTATCTTTATAGTCCCTGCGGTTCGCCACCC 3409
QY 5531 TCTGACTTGAGCGTTCGATTTTGTGATGCTGCTCAGGGGGCGGAGCCTATGGAAGAAACG 5590
DB 3410 TCTGACTTGAGCGTTCGATTTTGTGATGCTGCTCAGGGGGCGGAGCCTATGGAAGAAACG 3469
QY 5591 CGAGCAACGGCGCTTTTACGGTTCCTGCGCTTTTCTGCGCTTTTGTCTCACATGTCT 5650
DB 3470 CCAGCAACGGCGCTTTTACGGTTCCTGCGCTTTTCTGCGCTTTTGTCTCACATGTCT 3529
QY 5651 TTCTCGGTATCCCTGATTCTGTGTGATAACCGTATTACCGCTTTTGTAGTGTGAGTGATA 5710
DB 3530 TTCTCGGTATCCCTGATTCTGTGTGATAACCGTATTACCGCTTTTGTAGTGTGAGTGATA 3589
QY 5711 CCCTCGCCGAGCCGAAACGACCGAGCGCAGCGAGTCACTGAGCGAGGAAGCGGAAGCG 5770
DB 3590 CCGCTCGCCGAGCCGAAACGACCGAGCGCAGCGAGTCACTGAGCGAGGAAGCGGAAGCG 3649
QY 5771 GCCCAATACGCAACCGCTCTCCCGCGGCTTGGCCGATTCATTAATGCAG 5822
DB 3650 GCCCAATACGCAACCGCTCTCCCGCGGCTTGGCCGATTCATTAATGCAG 3701

RESULT 3
US-09-991-209-34
; Sequence 34, Application US/09991209
; Publication No. US20030024009A1
; GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Langdon, Timothy
; APPLICANT: Morse, Phillip
; TITLE OF INVENTION: Manipulation of the Phenolic Acid
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes
; FILE REFERENCE: GC648-2
; CURRENT APPLICATION NUMBER: US/09/991.209
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/249,608
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 4950
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pJ06.1 vector
US-09-991-209-34

Query Match 47.5%; Score 2768; DB 10; Length 4950;
Best local Similarity 96.8%; Pred. No. 0;
Matches 2895; Conservative 0; Mismatches 25; Indels 72; Gaps 4;

QY 2835 AAGCAGATCGTTCAACACATTTGGCAATAAAGTTTCTTAAGATTCGAATCCCTGTTGCCGGTC 2894
DB 955 AAGCAGATCGTTCAACACATTTGGCAATAAAGTTTCTTAAGATTCGAATCCCTGTTGCCGGTC 1014
QY 2895 TTGCGATGATTATCATATAATTTCTGTGTGAATTAAGTTAAGCATGTAATAAATTAACATGT 2954
DB 1015 TTGCGATGATTATCATATAATTTCTGTGTGAATTAAGTTAAGCATGTAATAAATTAACATGT 1074
QY 2955 AATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAAGTCCCGCAATTAATATTAATTT 3014
DB 1075 AATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAAGTCCCGCAATTAATATTAATTT 1134
QY 3015 AATACGGGATGAGAAACAAATATAGCGCGCAACTAGGATTAATTAATTCGGCGCGGTGT 3074
```

```
DB 1135 AATAACGGATAGAAAACAAATATATAGCGCGCAAACTAGGATTAATTAATTCGGCGCGGTGT 1194
QY 3075 CATCTATGTTACTAGATTCGACCTTCGAGGCAATGGAATCCGGCGCGCAATGCAATCGCGGC 3134
DB 1195 CATCTATGTTACTAGATTCGATAAGCTTCTAGAGCGCGCGGTGGAGC-----T 1241
QY 3135 CCAATTCGCCCTATAGTGTCTGTATTAACCGCGGCTCCTAGCGCGTCTGTATTAACAAGTC 3191
DB 1242 CCAATTCGCCCTATAGTGTCTGTATTAACCGCGGCTCCTAGCGCGTCTGTATTAACAAGTC 1301
QY 3192 GTGACTTGGGAAAAACCTTGGCGTTTACCAACTTAATTCGCTTTCGAGCAATCCCTTTCG 3251
DB 1302 GTGACTTGGGAAAAACCTTGGCGTTTACCAACTTAATTCGCTTTCGAGCAATCCCTTTCG 1361
QY 3252 CGAGCTTGGCGTAATAGCGAAGAGCGCGCAACCGATTCGCTTTCGCAACAGTTGGCGAGCC 3311
DB 1362 CGAGCTTGGCGTAATAGCGAAGAGCGCGCAACCGATTCGCTTTCGCAACAGTTGGCGAGCC 1421
QY 3312 TGAATGGCGAAT--GGAGCGCCCTGTAGCGCGCATTTAAGCGCGGCGGGTGTGGTGA 3370
DB 1422 TGAATGGCGAATGGAGCGCCCTGTAGCGCGCATTTAAGCGCGGCGGGTGTGGTGA 1481
QY 3371 CGCGCAGCGTGAACCGCTTACACTTTCGCAAGCGCGCTTTCGCTTTCCTTTC 3430
DB 1482 CGCGCAGCGTGAACCGCTTACACTTTCGCAAGCGCGCTTTCGCTTTCCTTTC 1541
QY 3431 CTTCTTTTCTCGCACGTTTCGCGGCTTTCGCGCTCAAGCTCTTAATTCGGGGGCTCCCTT 3490
DB 1542 CTTCTTTTCTCGCACGTTTCGCGGCTTTCGCGCTCAAGCTCTTAATTCGGGGGCTCCCTT 1601
QY 3491 TAGGGTTCGGATTTAGAGCTTTACGGCACTTCGACCGCAAAACCTTGAATTTGGTGTG 3550
DB 1602 TAGGGTTCGGATTTAGAGCTTTACGGCACTTCGACCGCAAAACCTTGAATTTGGTGTG 1661
QY 3551 GTTCACGCTAGTGGGCCATCGCCCTGTAGAGCGGTTTTTCGCCCTTTCGAGTCCA 3610
DB 1662 GTTCACGCTAGTGGGCCATCGCCCTGTAGAGCGGTTTTTCGCCCTTTCGAGTCCA 1721
QY 3611 CGTCTTTAATAGTGGACTCTTGTTCAAACTGGAACAACTCAACCTATCTCGGTCT 3670
DB 1722 CGTCTTTAATAGTGGACTCTTGTTCAAACTGGAACAACTCAACCTATCTCGGTCT 1781
QY 3671 ATTCTTTTGTATTTATAGGGAATTTGCGGCTATTCGGCTTAAAAAATCAGCTGA 3730
DB 1782 ATTCTTTTGTATTTAATAGGGAATTTGCGGCTATTCGGCTTAAAAAATCAGCTGA 1841
QY 3731 TTTAAACAAATATTTAAACCGCAATTTTAAACAAATATTTAAACCTTTGCGCTGATG 3790
DB 1842 TTTAAACAAATATTTAAACCGCAATTTTAAACAAATATTTAAACCTTTGCGCTGATG 1892
QY 3791 CGGTATTTCTCTTACGCATCTGTGCGGTATTTTCAACCGCATACAGTGGCCTTTTC 3850
DB 1893 -----AGTGGCCTTTTTC 1906
QY 3851 GGGGAATGTCGGGAAACCCCTATTTGTTTATTTTCTTAATATACATTTCAATATGATC 3910
DB 1907 GGGGAATGTCGGGAAACCCCTATTTGTTTATTTTCTTAATATACATTTCAATATGATC 1966
QY 3911 CGCTCATGAGCAATAACCTGATAAATGCTTCAATAATATTGAAAAAGGAAGATGA 3970
DB 1967 CGCTCATGAGCAATAACCTGATAAATGCTTCAATAATATTGAAAAAGGAAGATGA 2026
QY 3971 GTATTCAAATTTCCGTGTCGCCCTTATTCCTTTTTTGGGCAATTTTGCCTTCTGTTT 4030
DB 2027 GTATTCAAATTTCCGTGTCGCCCTTATTCCTTTTTTGGGCAATTTTGCCTTCTGTTT 2086
QY 4031 TTGCTCAACCCAGAAACCTGCTGAACTAAAGATGCTGAAGATCAGTTGGTGCACGAG 4090
DB 2087 TTGCTCAACCCAGAAACCTGCTGAACTAAAGATGCTGAAGATCAGTTGGTGCACGAG 2146
QY 4091 TGGTTACATCGAATCTGGATCTCAACAGCGGTAAAGATCTTTAGAGATTTTTCGCCCGAAG 4150
DB 2147 TGGTTACATCGAATCTGGATCTCAACAGCGGTAAAGATCTTTAGAGATTTTTCGCCCGAAG 2206
```

```
QY 4151 AACGTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATATGTGGCGGGTATTAATCCCGTA 4210
Db 2207 AACGTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATATGTGGCGGGTATTAATCCCGTA 2266
QY 4211 TTGACGCCGGGCAAGAGCAACTCGGTGCGCGCATACACATATCTCAGAAATGACTTGGTGG 4270
Db 2267 TTGACGCCGGGCAAGAGCAACTCGGTGCGCGCATACACATATCTCAGAAATGACTTGGTGG 2326
QY 4271 AGTACTCACAGTCACAGAAAGCACTTAACGATGGCATGACAGTAAGAGAAATATGCA 4330
Db 2327 AGTACTCACAGTCACAGAAAGCACTTAACGATGGCATGACAGTAAGAGAAATATGCA 2386
QY 4331 GTGCTGCATAACCAATGATGATAAAGTCTGGGCGCAACTTACTCTGACAAAGATCGGAG 4390
Db 2387 GTGCTGCATAACCAATGATGATAAAGTCTGGGCGCAACTTACTCTGACAAAGATCGGAG 2446
QY 4391 GACCGAAGGAGCTAAACCGCTTTTTCACAAACATGGGGGATCATGTAACCTGGCTTCATC 4450
Db 2447 GACCGAAGGAGCTAAACCGCTTTTTCACAAACATGGGGGATCATGTAACCTGGCTTCATC 2506
QY 4451 GTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGAACACCAAGATGCTG 4510
Db 2507 GTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGAACACCAAGATGCTG 2566
QY 4511 TAGCAATGGCAACAAACGTTGCGCAAACTATTAACCTGGCGAACTACTTACTAGCTTCCC 4570
Db 2567 TAGCAATGGCAACAAACGTTGCGCAAACTATTAACCTGGCGAACTACTTACTAGCTTCCC 2626
QY 4571 GGCACAAATTAATAGCTGATGAGCGGATGAAGTTGCGAGCACTTCTGGCTCGG 4630
Db 2627 GGCACAAATTAATAGCTGATGAGCGGATGAAGTTGCGAGCACTTCTGGCTCGG 2686
QY 4631 CCCTTCGGGCTGGCTGTTATGCTGATAAATCTGGAGCGGTGAGCGTGGGTCTCGCG 4690
Db 2687 CCCTTCGGGCTGGCTGTTATGCTGATAAATCTGGAGCGGTGAGCGTGGGTCTCGCG 2746
QY 4691 GTATCATTTGACGACTGGGGCGAGATGGTAAGCCCTCCCGTATGTAAGTTATCTACAGA 4750
Db 2747 GTATCATTTGACGACTGGGGCGAGATGGTAAGCCCTCCCGTATGTAAGTTATCTACAGA 2806
QY 4751 CGGGGAGTCAGGCAACTATGATGAACGAATAGACAGATCGCTGAGATGGTCCCTCAC 4810
Db 2807 CGGGGAGTCAGGCAACTATGATGAACGAATAGACAGATCGCTGAGATGGTCCCTCAC 2866
QY 4811 TGATTAAGCAATGGTAAGTCTGACACCAAGTTTACTCATATATCTTATAGATGATTA 4870
Db 2867 TGATTAAGCAATGGTAAGTCTGACACCAAGTTTACTCATATATCTTATAGATGATTA 2926
QY 4871 AACTTCATTTTTTAAATAAGGATCTAGGTGAAGATCCCTTTTGTGATAATCTCATGACCA 4930
Db 2927 AACTTCATTTTTTAAATAAGGATCTAGGTGAAGATCCCTTTTGTGATAATCTCATGACCA 2986
QY 4931 AATCCCTTAACGTAGTCTTCTGTCACATGAGCGTCAAGCCCGGTAGAAAGATCAAG 4990
Db 2987 AATCCCTTAACGTAGTCTTCTGTCACATGAGCGTCAAGCCCGGTAGAAAGATCAAG 3046
QY 4991 GATCTCTCTGAGATCCCTTTTCTGCGCGTAATCTGCTGCTGTCGCAACAAAAAACCC 5050
Db 3047 GATCTCTCTGAGATCCCTTTTCTGCGCGTAATCTGCTGCTGTCGCAACAAAAAACCC 3106
QY 5051 CGCTACAGCGGTGTTTGTGTCGAGATCAAGAGCTACCACTCTTTTCCGAAAGTAA 5110
Db 3107 CGCTACAGCGGTGTTTGTGTCGAGATCAAGAGCTACCACTCTTTTCCGAAAGTAA 3166
QY 5111 CTGCTCTCAGCAGCGGATACCAATCTGCTCTAGTGTAGCGGTAGTGTAGGCC 5170
Db 3167 CTGCTCTCAGCAGCGGATACCAATCTGCTCTAGTGTAGCGGTAGTGTAGGCC 3226
QY 5171 ACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCTCTGCTAATCTGTTAGCAG 5230
Db 3227 ACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCTCTGCTAATCTGTTAGCAG 3286
```

```
QY 5231 TGCTGTGCTCCAGTGGCGATAAGTCTGCTTACCGGGTTGGACTCAGACGATAGTTAC 5290
Db 3287 TGCTGTGCTCCAGTGGCGATAAGTCTGCTTACCGGGTTGGACTCAGACGATAGTTAC 3346
QY 5291 CGGATAAGCGCGACGCGTCCGGCTGAACGGGGGTTCTGTGACACAGCCAGCTTGGAGC 5350
Db 3347 CGGATAAGCGCGACGCGTCCGGCTGAACGGGGGTTCTGTGACACAGCCAGCTTGGAGC 3406
QY 5351 GAACGACTTACACCGAACTGAGATACCTACAGCTGTAGCTATGAGAAAGCGCCAGCTTC 5410
Db 3407 GAACGACTTACACCGAACTGAGATACCTACAGCTGTAGCTATGAGAAAGCGCCAGCTTC 3466
QY 5411 CCGAAGGAGAAAGCGGACAGCTATCCGTGAAGCGGAGCTGGACAGGAGCGCA 5470
Db 3467 CCGAAGGAGAAAGCGGACAGCTATCCGTGAAGCGGAGCTGGACAGGAGCGCA 3526
QY 5471 CGAGGAGCTTCCAGGGGCAAAACCGCTGGTATCTTTATAGTCTCTGTGCGGTTTCCGCCACC 5530
Db 3527 CGAGGAGCTTCCAGGGGCAAAACCGCTGGTATCTTTATAGTCTCTGTGCGGTTTCCGCCACC 3586
QY 5531 TCTGACTTGAAGCGTTCGATTTTGTGATGCTGCTCAGGGGGCGGAGCCTATGGAACAAACG 5590
Db 3587 TCTGACTTGAAGCGTTCGATTTTGTGATGCTGCTCAGGGGGCGGAGCCTATGGAACAAACG 3646
QY 5591 CCAGCAACCGGCTTTTAAAGGTTCTGCGCTTTTCTGCGCTTTTCTGCTCAGATGTTCT 5650
Db 3647 CCAGCAACCGGCTTTTAAAGGTTCTGCGCTTTTCTGCGCTTTTCTGCTCAGATGTTCT 3706
QY 5651 TTCTGCGTTTATCCCTGATTTCTGTGATTAACCGTATTTACCGCTTTTGTAGTGAGCTGATA 5710
Db 3707 TTCTGCGTTTATCCCTGATTTCTGTGATTAACCGTATTTACCGCTTTTGTAGTGAGCTGATA 3766
QY 5711 CCGCTGCCCGCAGCGCAACCGACGCGCAGCTGAGTCACTGAGCGAGGAGCGGAGAGC 5770
Db 3767 CCGCTGCCCGCAGCGCAACCGACGCGCAGCTGAGTCACTGAGCGAGGAGCGGAGAGC 3826
QY 5771 GCCCAATACGCAACCGCTCTCCCGCGGTTGGCGGATTCATTTAATGCGAG 5822
Db 3827 GCCCAATACGCAACCGCTCTCCCGCGGTTGGCGGATTCATTTAATGCGAG 3878

RESULT 4
US-09-991-209-37
; Sequence 37, Application US/09991209
; Publication No. US20030024009A1
; GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Langdon, Timothy
; APPLICANT: Morse, Phillip
; TITLE OF INVENTION: Manipulation of the Phenolic Acid
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes
; FILE REFERENCE: GC648-2
; CURRENT APPLICATION NUMBER: US/09/991.209
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/249,608
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 4965
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pJQ3 vector
US-09-991-209-37

Query Match 47.5%; Score 2768; DB 10; Length 4965;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 2895; Conservative 0; Mismatches 25; Indels 72; Gaps 4;

QY 2835 AAGCAGATCGTTCAAAACATTGGCAATAAAGTTTCTTAAGATTCGAATCTGTTGCGGTC 2894
|||||
```

Db 1824 AAGCAGATCGTTCAACCAATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTC 1883  
Qy 2895 TTGGGATGATATCATATAATTTCTGTGTGAATTAAGTTAGCATGTAATTAATAACATGT 2954  
Db 1884 TTGGGATGATATCATATAATTTCTGTGTGAATTAAGTTAGCATGTAATTAATAACATGT 1943  
Qy 2955 AATGCAATGATTAATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTAACATTT 3014  
Db 1944 AATGCAATGATTAATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTAACATTT 2003  
Qy 3015 AATACCGGATAGAAAACAAATAATAGCGCAAACTAGGATAAATTAATCGCGCGGTGT 3074  
Db 2004 AATACCGGATAGAAAACAAATAATAGCGCAAACTAGGATAAATTAATCGCGCGGTGT 2063  
Qy 3075 CATCTATGTTACTAGATCGACCTGCGAGGATGAGATCCGCGCGGATGCGAGTGGGC 3134  
Db 2064 CATCTATGTTACTAGATCGAATAGCTTCTAGAGCGCGCGGTGGAGC-----T 2110  
Qy 3135 CCAATTGCGCCTATAGTAGAGTCGTATTAGC--AATTCACTGGCGCTGTTTATCAACGTC 3191  
Db 2111 CCAATTGCGCCTATAGTAGAGTCGTATTAGCGGCGCTCACTGGCGCTGTTTATCAACGTC 2170  
Qy 3192 GTGACTGGGAAAACCCCTGGGTTTACCCAACTTAATGCGCTTTCAGGACATCCCGCTTTCG 3251  
Db 2171 GTGACTGGGAAAACCCCTGGGTTTACCCAACTTAATGCGCTTTCAGGACATCCCGCTTTCG 2230  
Qy 3252 CCAAGTGGCGTAATAGCGAAGAGGCGCGCAACCGATGCGCTTTCAGGACATCCCGCTTTCG 3311  
Db 2231 CCAAGTGGCGTAATAGCGAAGAGGCGCGCAACCGATGCGCTTTCAGGACATCCCGCTTTCG 2290  
Qy 3312 TGAATGCGCAAT--GGAGCGGCCCTGATAGCGCGCATTAAGCGCGGGGTGCGTGTGA 3370  
Db 2291 TGAATGCGCAATGGAGCGGCCCTGATAGCGCGCATTAAGCGCGGGGTGCGTGTGA 2350  
Qy 3371 CGCGCAGCTGACCGCTTACACTTTCAGCGCGCCCTAGCGCGCGCTTTCGCTTTCCTTC 3430  
Db 2351 CGCGCAGCTGACCGCTTACACTTTCAGCGCGCCCTAGCGCGCGCTTTCGCTTTCCTTC 2410  
Qy 3431 CTTTCCTTTTCGCGCAGCTTTCGCGGCTTTCGCGTCAAGCTCTAAATCGGGGCTTCCTT 3490  
Db 2411 CTTTCCTTTTCGCGCAGCTTTCGCGGCTTTCGCGTCAAGCTCTAAATCGGGGCTTCCTT 2470  
Qy 3491 TAGGGTTCCGATTTAGAGCTTTACGGCACCTCGACCGCAAAACTTTGATTGGGTGATG 3550  
Db 2471 TAGGGTTCCGATTTAGTCTTTAGCGCACCTCGACCGCAAAACTTTGATTAGGGTGTG 2530  
Qy 3551 GTTCAGCTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCA 3610  
Db 2531 GTTCAGCTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCA 2590  
Qy 3611 CGTTCTTTAATAGTGGACTCTGTTCCAACTGGACCAACACTCAACCCCTATCTCGTCT 3670  
Db 2591 CGTTCTTTAATAGTGGACTCTGTTCCAACTGGACCAACACTCAACCCCTATCTCGTCT 2650  
Qy 3671 ATTCTTTTGAATTTAAGGATTTTCGGATTTTCGGCTATTCGTTTAAATAATAGCTGA 3730  
Db 2651 ATTCTTTTGAATTTAAGGATTTTCGGATTTTCGGCTATTCGTTTAAATAATAGCTGA 2710  
Qy 3731 TTTTAAACAAATTAATTAACGCGAAATTTTAAACAAATATTAACGTTTACAAATTTTCGCTGATG 3790  
Db 2711 TTTTAAACAAATTTTAAACGCGAAATTTTAAACAAATATTAACGTTTACAAATTT----- 2761  
Qy 3791 CGGTATTTCTCCTTACGCACTGTCGGGTATTTCACACCGCATACAGGTGGCACTTTTC 3850  
Db 2762 -----AGGTGGCACTTTTC 2775  
Qy 3851 GGGGAAATGTCGCGGAACCCCTATTGTTTATTTTCTAAATACATTTCAAAATATGATC 3910  
Db 2776 GGGGAAATGTCGCGGAACCCCTATTGTTTATTTTCTAAATACATTTCAAAATATGATC 2835  
Qy 3911 CGCTCATGAGCAATAACCCCTGATAATGCTTCAATAATTTGAAAGAGGAGATGA 3970  
Db 2836 CGCTCATGAGCAATAACCCCTGATAATGCTTCAATAATTTGAAAGAGGAGATGA 2895

Qy 3971 GTATTCAACATTTCCGTTGCGCCTTAATTCCTCTTTTGGGCAATTTTGCCTTCTCTGTTT 4030  
Db 2896 GTATTCAACATTTCCGTTGCGCCTTAATTCCTCTTTTGGGCAATTTTGCCTTCTCTGTTT 2955  
Qy 4031 TTGCTCACCCAGAAAACGCTGCTGAAGTAAAGATGCTGAAGATCAGTTGGGTGCAACGAG 4090  
Db 2956 TTGCTCACCCAGAAAACGCTGCTGAAGTAAAGATGCTGAAGATCAGTTGGGTGCAACGAG 3015  
Qy 4091 TGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAAGATTTTCGCCCCGAG 4150  
Db 3016 TGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAAGATTTTCGCCCCGAG 3075  
Qy 4151 AAGTTTTTCAATGATGAGCACTTTTAAAGTTTCTGCTATGTTGGCGGCTATTATCCCGTA 4210  
Db 3076 AAGTTTTTCAATGATGAGCACTTTTAAAGTTTCTGCTATGTTGGCGGCTATTATCCCGTA 3135  
Qy 4211 TTCAACGCGGCAAGAGCAACTCGTCCGCGCATACACTATTCTCAGAATGACTTGGTTG 4270  
Db 3136 TTCAACGCGGCAAGAGCAACTCGTCCGCGCATACACTATTCTCAGAATGACTTGGTTG 3195  
Qy 4271 AGTACTCACCGTCAAGAAAAGCATCTTACGCGATGGCATGACAGTAAGAGAAATATGCA 4330  
Db 3196 AGTACTCACCGTCAAGAAAAGCATCTTACGCGATGGCATGACAGTAAGAGAAATATGCA 3255  
Qy 4331 GTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTCTCACAACGATCGGAG 4390  
Db 3256 GTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTCTCACAACGATCGGAG 3315  
Qy 4391 GACCGAAGAGAGCTTAAACCGCTTTTTCACAACATGCGGGATCATGTAACTCGCCTTGATC 4450  
Db 3316 GACCGAAGAGAGCTTAAACCGCTTTTTCACAACATGCGGGATCATGTAACTCGCCTTGATC 3375  
Qy 4451 GTTGGAAACCGGAGCTGAATGAAGCCATACCAAACGAGCGTGAACACCAAGATGCGCTG 4510  
Db 3376 GTTGGAAACCGGAGCTGAATGAAGCCATACCAAACGAGCGTGAACACCAAGATGCGCTG 3435  
Qy 4511 TAGCAATGGCAACAAAGCTTTCGCAAACTATTAACTGCGCAACTTACTCTAGCTTCCC 4570  
Db 3436 TAGCAATGGCAACAAAGCTTTCGCAAACTATTAACTGCGCAACTTACTCTAGCTTCCC 3495  
Qy 4571 GGCAACAAATTAATAGATCGGATGGAGCGGATAAAGTTGCGAGGACCACTTCTGCGCTCG 4630  
Db 3496 GGCAACAAATTAATAGATCGGATGGAGCGGATAAAGTTGCGAGGACCACTTCTGCGCTCG 3555  
Qy 4631 CCCTTCGCGCTGCGTGTGTTTATTCGTGATTAATCTGAGCGCGTGAAGCTGCGTCCG 4690  
Db 3556 CCCTTCGCGCTGCGTGTGTTTATTCGTGATTAATCTGAGCGCGTGAAGCTGCGTCCG 3615  
Qy 4691 GTATCATTTGAGCACTGGGCGCAGATGGTAAGCCCTCCGCTATCTAGTTATCTACACGA 4750  
Db 3616 GTATCATTTGAGCACTGGGCGCAGATGGTAAGCCCTCCGCTATCTAGTTATCTACACGA 3675  
Qy 4751 CGGGAGTTCAGGCAACTATCGATGAACGAAATAGACAGATCGCTGAGATAGGTGCTCAC 4810  
Db 3676 CGGGAGTTCAGGCAACTATCGATGAACGAAATAGACAGATCGCTGAGATAGGTGCTCAC 3735  
Qy 4811 TGAATTAAGCAATGGTTAACTGTCAGACCAAGTTTACTCATATATATCTTTAGATTGATTAA 4870  
Db 3736 TGAATTAAGCAATGGTTAACTGTCAGACCAAGTTTACTCATATATATCTTTAGATTGATTAA 3795  
Qy 4871 AACTTCATTTTAAATTTAAAGGATCTAGTGAAGATCCTTTTGTATTAATCTCATGACCA 4930  
Db 3796 AACTTCATTTTAAATTTAAAGGATCTAGTGAAGATCCTTTTGTATTAATCTCATGACCA 3855  
Qy 4931 AAATCCCTTAAACGTTAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAG 4990  
Db 3856 AAATCCCTTAAACGTTAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAG 3915  
Qy 4991 GATCTCTTGAGATCCTTTTCTGCGCGTAACTGCTGCTTGCCTGCAAAACAAAACACAC 5050  
Db 3916 GATCTCTTGAGATCCTTTTCTGCGCGTAACTGCTGCTTGCCTGCAAAACAAAACACAC 3975





QY 3791 CGGTATTTTCTCCTTACGCACTCTGCGCGGTATTTTCAACCGCATACAGGTGGCACTTTTC 3850  
DB 2771 -----AGGTGGCACTTTTC 2784  
QY 3851 GGGGAATGTGCGGGAACCCCTATTATTTGTTTATTTTCTAAATACATTTCAAAATATGATC 3910  
DB 2785 GGGGAATGTGCGGGAACCCCTATTATTTGTTTATTTTCTAAATACATTTCAAAATATGATC 2844  
QY 3911 CGCTCATGAGACAATAAACCTCTGATAAATGCTTCAATATATTTGAAAAAGGAAGATGA 3970  
DB 2845 CGCTCATGAGACAATAAACCTCTGATAAATGCTTCAATATATTTGAAAAAGGAAGATGA 2904  
QY 3971 GTATTCAACATTTCCGCTGTCGCCCTTATTCCTTTTGTGGGCATTTTGCCTTCTGTTT 4030  
DB 2905 GTATTCAACATTTCCGCTGTCGCCCTTATTCCTTTTGTGGGCATTTTGCCTTCTGTTT 2964  
QY 4031 TTGCTCACCCAGAAACCTGCTGTAAGTAAAGATGCTGAAGATCAGTTCGGTGCAAG 4090  
DB 2965 TTGCTCACCCAGAAACCTGCTGTAAGTAAAGATGCTGAAGATCAGTTCGGTGCAAG 3024  
QY 4091 TGGGTTACATCGAATCGAATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCGAAG 4150  
DB 3025 TGGGTTACATCGAATCGAATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCGAAG 3084  
QY 4151 AACGTTTTCGAATGATGAGCACTTTTAAAGTTCTGCTATGCTGCGCGGTATTTACCGTA 4210  
DB 3085 AACGTTTTCGAATGATGAGCACTTTTAAAGTTCTGCTATGCTGCGCGGTATTTACCGTA 3144  
QY 4211 TTGACGCGGGCAAGAGCACTCGTTCGCCGATACACTATTCTCAGAATGACTTGGTTC 4270  
DB 3145 TTGACGCGGGCAAGAGCACTCGTTCGCCGATACACTATTCTCAGAATGACTTGGTTC 3204  
QY 4271 AGTACTCACGATCACAGAAAAGCATCTTAACGATGCGATGACAGTAAGAAATATGCA 4330  
DB 3205 AGTACTCACGATCACAGAAAAGCATCTTAACGATGCGATGACAGTAAGAAATATGCA 3264  
QY 4331 GTCTCTGCCATAACCATGAGTATACACTCGCGGCACTTACTTCTGACAGAACATCGGAG 4390  
DB 3265 GTCTCTGCCATAACCATGAGTATACACTCGCGGCACTTACTTCTGACAGAACATCGGAG 3324  
QY 4391 GACCGAAGAGCTAAACCGCTTTTGTCAACAATGCGGATCATGTAACTCGCTTGATC 4450  
DB 3325 GACCGAAGAGCTAAACCGCTTTTGTCAACAATGCGGATCATGTAACTCGCTTGATC 3384  
QY 4451 GTTGGAAACCGGAGCTGAATGAAGCCATACCAAACGACGCTGACACCAAGATGCTG 4510  
DB 3385 GTTGGAAACCGGAGCTGAATGAAGCCATACCAAACGACGCTGACACCAAGATGCTG 3444  
QY 4511 TAGCAATGGCAACACGTTGCGCAACTATTAACTGCGCACTACTTACTCTAGCTTCCC 4570  
DB 3445 TAGCAATGGCAACACGTTGCGCAACTATTAACTGCGCACTACTTACTCTAGCTTCCC 3504  
QY 4571 GGCACAAATTAATAGACTGGAGCGGATAAAGTTTGACGACCACTTCTCGCTCGG 4630  
DB 3505 GGCACAAATTAATAGACTGGAGCGGATAAAGTTTGACGACCACTTCTCGCTCGG 3564  
QY 4631 CCCTTCGGCTGCTGTTTATTTGCTGATAAATCTGAGACCGGTGAGCGGTCTCGG 4690  
DB 3565 CCCTTCGGCTGCTGTTTATTTGCTGATAAATCTGAGACCGGTGAGCGGTCTCGG 3624  
QY 4691 GTATCATTTGACGACTTGGGCGGATGTTAAGCCCTCCGATCGTAGTTATCTACA 4750  
DB 3625 GTATCATTTGACGACTTGGGCGGATGTTAAGCCCTCCGATCGTAGTTATCTACA 3684  
QY 4751 CCGGGAGTCAGGCAACTATGATGAACGAATAGACAGATCGCTGAGATAGGTGCTCAC 4810  
DB 3685 CCGGGAGTCAGGCAACTATGATGAACGAATAGACAGATCGCTGAGATAGGTGCTCAC 3744  
QY 4811 TGAATTAAGCATTTGGTAACTGTGACACCAAGTTTACTCATATATCTTTAGATTTAA 4870  
DB 3745 TGAATTAAGCATTTGGTAACTGTGACACCAAGTTTACTCATATATCTTTAGATTTAA 3804

QY 4871 AACTTCATTTTAAATTTAAAGGATCTAGTGAAGATCCCTTTTGTATATCTCATGACCA 4930  
DB 3805 AACTTCATTTTAAATTTAAAGGATCTAGTGAAGATCCCTTTTGTATATCTCATGACCA 3864  
QY 4931 AAATCCCTTAAACGTAGCTTTTGTTCACCTGAGCGTCAAGCCCGTAGAAAAGATCAAG 4990  
DB 3865 AAATCCCTTAAACGTAGCTTTTGTTCACCTGAGCGTCAAGCCCGTAGAAAAGATCAAG 3924  
QY 4991 GATCTTCTGAGATCCCTTTTCTGCGGTAAATCTGCTGCTTGCGAAACAAAAAACCAAC 5050  
DB 3925 GATCTTCTGAGATCCCTTTTCTGCGGTAAATCTGCTGCTTGCGAAACAAAAAACCAAC 3984  
QY 5051 CGCTTACAGCGGTGTTTGTGCGGATCAAGAGCTACCAACTCTTTTCCGAGGTAA 5110  
DB 3985 CGCTTACAGCGGTGTTTGTGCGGATCAAGAGCTACCAACTCTTTTCCGAGGTAA 4044  
QY 5111 CTGCTTTCAGCAGCGCAGATACAAATACCTGCTCTTAGTGTAGCCGTAGTTAGGCC 5170  
DB 4045 CTGCTTTCAGCAGCGCAGATACCAATACCTGCTCTTAGTGTAGCCGTAGTTAGGCC 4104  
QY 5171 ACCACTTCAAGAACTCTGTAGCACCGCTCATACCTGCTCTGCTAAATCTGTTACCAG 5230  
DB 4105 ACCACTTCAAGAACTCTGTAGCACCGCTCATACCTGCTCTGCTAAATCTGTTACCAG 4164  
QY 5231 TGGCTGCTGCGAGTGGCGATAAGTCTTACCGGGTGGACTCAAGACGATAGTTAC 5290  
DB 4165 TGGCTGCTGCGAGTGGCGATAAGTCTTACCGGGTGGACTCAAGACGATAGTTAC 4224  
QY 5291 CGGATAAGCGCGCAGCGCTCGGCTGAACCGGGGTCTGTCACACAGCCAGCTTGGAGC 5350  
DB 4225 CGGATAAGCGCGCAGCGCTCGGCTGAACCGGGGTCTGTCACACAGCCAGCTTGGAGC 4284  
QY 5351 GAACGACTTACACCGAATCAGATACCTACAGCTGAGCTATGAGAAAGCCGCTTC 5410  
DB 4285 GAACGACTTACACCGAATCAGATACCTACAGCTGAGCTATGAGAAAGCCGCTTC 4344  
QY 5411 CCGAAGGAGAAAGCGGACAGGTATCCGCTAGCGGCGAGGTCGGAACAGAGAGCGCA 5470  
DB 4345 CCGAAGGAGAAAGCGGACAGGTATCCGCTAGCGGCGAGGTCGGAACAGAGAGCGCA 4404  
QY 5471 CCGAGGAGCTTCCAGGGGAAACCGCTGCTATCTTATAGTCTGCTCGGGTTCCGCCAC 5530  
DB 4405 CCGAGGAGCTTCCAGGGGAAACCGCTGCTATCTTATAGTCTGCTCGGGTTCCGCCAC 4464  
QY 5531 TCTGACTTGAGCGTCTGATCTGCTCAGGGGGCGGAGCTATGGAANAACG 5590  
DB 4465 TCTGACTTGAGCGTCTGATCTGCTCAGGGGGCGGAGCTATGGAANAACG 4524  
QY 5591 CCAGCAACGGCGCTTTTACGCTTCTGCGCTTTTGTGCTGCTGCTGCTGCTGCTGCT 5650  
DB 4525 CCAGCAACGGCGCTTTTACGCTTCTGCGCTTTTGTGCTGCTGCTGCTGCTGCTGCT 4584  
QY 5651 TTCTGCTTATCCCTGATTTCTGAGTAAACGCTATTTACCGCTTTGAGTGAAGCTGATA 5710  
DB 4585 TTCTGCTTATCCCTGATTTCTGAGTAAACGCTATTTACCGCTTTGAGTGAAGCTGATA 4644  
QY 5711 CCCTGCGCGCAGCGCAACGACCGAGCGCAGCTCAGTGAAGCGAGCGGAGCGAGC 5770  
DB 4645 CCCTGCGCGCAGCGCAACGACCGAGCGCAGCTCAGTGAAGCGAGCGGAGCGAGC 4704  
QY 5771 GCCCAATAGCAAAACCGCTTCTCCCGCGCTTGGCGGATTCATTAATGCA 5822  
DB 4705 GCCCAATAGCAAAACCGCTTCTCCCGCGCTTGGCGGATTCATTAATGCA 4756

## RESULT 6

US-09-991-209-40  
; Sequence 40, Application US/09991209  
; Publication No. US20030024009A1  
; GENERAL INFORMATION:  
; APPLICANT: Dunn-Coleman, Nigel  
; APPLICANT: Langdon, Timothy  
; APPLICANT: Morse, Phillip

;  
; TITLE OF INVENTION: Manipulation of the Phenolic Acid  
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted  
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes  
; FILE REFERENCE: GC648-2  
; CURRENT APPLICATION NUMBER: US/09/991,209  
; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: US 60/249,608  
; PRIOR FILING DATE: 2000-11-17  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 5001  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pub8.11 vector  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(5001)  
; OTHER INFORMATION: n = A,T,C or G  
; US-09-991-209-40

Query Match 47.5%; Score 2768; DB 10; Length 5001;  
Best Local Similarity 96.8%; Pred. No. 0;  
Matches 2895; Conservative 0; Mismatches 25; Indels 72; Gaps 4;

Qy 2835 AAGCAGATCGTTCAAAACATTTGGCAATAAAGTTTCTTAAGATTTGAATCCTGTGCGGTC 2894  
Db |||||  
Qy 2895 TTGGATGATTTATCATATAATTTCTGTGAATAGTTAGTTAAGCATGTAAATTAACATGT 2954  
Db |||||  
Qy 1916 TTGGATGATTTATCATATAATTTCTGTGTGAATAGTTAAGCATGTAAATTAACATGT 1975  
Db |||||  
Qy 2955 AATCATGACGTTATTTATGAGATGGTTTATGATTTAGATTCGCGCAATTTATACATTT 3014  
Db |||||  
Qy 1976 AATCATGACGTTATTTATGAGATGGTTTATGATTTAGATTCGCGCAATTTATACATTT 2035  
Db |||||  
Qy 3015 AATACGGATAGAAACAAATATAGCGCGCAAACTAGGATAAAATTTATCGCGCGGTGT 3074  
Db |||||  
Qy 2036 AATACGGATAGAAACAAATATAGCGCGCAAACTAGGATAAAATTTATCGCGCGGTGT 2095  
Db |||||  
Qy 3075 CATCTATGTTACTAGATCGACTCGAGGCATGGATCGCGCGCGCATGCGACGTGCGGC 3134  
Db |||||  
Qy 2096 CATCTATGTTACTAGATCGATTAAGCTTCTAGAGCGCGGTGGAGC-----T 2142  
Db |||||  
Qy 3135 CCAATTCGCGCTATAGTGAAGTCTGATTAC---AATTCACTGCGCGTGTGTTTACAGCTC 3191  
Db |||||  
Qy 2143 CCAATTCGCGCTATAGTGAAGTCTGATTAGCGCGGTCTACTGCGCGTCTGTTTACAGCTC 2202  
Db |||||  
Qy 3192 GTGACTGGGAAACCTTGGCGTTACCCAACTTAATGCGCTTTCGACACATCCCCCTTTTCG 3251  
Db |||||  
Qy 2203 GTGACTGGGAAACCTTGGCGTTTACCCAACTTAATGCGCTTTCGACACATCCCCCTTTTCG 2262  
Db |||||  
Qy 3252 CCAGCTGGCGTAATAGCGAGAGCGCCGACCGATCGCCCTTCCACAGTTGCGGACGC 3311  
Db |||||  
Qy 2263 CCAGCTGGCGTAATAGCGAGAGCGCCGACCGATCGCCCTTCCACAGTTGCGGACGC 2322  
Db |||||  
Qy 3312 TGAATGCGGAAT-GGACGCGCTGTAGCGCGGATTAAGCGCGCGGTGTGTTGTTA 3370  
Db |||||  
Qy 2323 TGAATGCGGAATGGAAGCGCGCTGTAGCGCGGATTAAGCGCGCGGTGTGTTGTTA 2382  
Db |||||  
Qy 3371 CGCGAGCGGTGACCGCTACACTTTCGACGCGCTTAGCGCGCGCTTTCGCTTTCTTTC 3430  
Db |||||  
Qy 2383 CGCGAGCGGTGACCGCTACACTTTCGACGCGCTTAGCGCGCGCTTTCGCTTTCTTTC 2442  
Db |||||  
Qy 3431 CTTCTTTCTCGCCACGTTTCGCGCGCTTTTCGCGCTGAACTCTAAATCGCGGCGCTTCCCTT 3490  
Db |||||  
Qy 2443 CTTCTTTCTCGCCACGTTTCGCGCGCTTTTCGCGCTGAACTCTAAATCGCGGCGCTTCCCTT 2502  
Db |||||  
Qy 3491 TAGGGTTCCGATTTAGAGCTTTAGGCACTTCGACCGCAAAACTTGAATTTGGGTGATG 3550  
Db |||||  
Qy 2503 TAGGGTTCCGATTTAGTGTCTTTAGGCACTTCGACCGCAAAACTTGAATTTAGGGTGTATG 2562  
Db |||||

Qy 3551 GTTCAAGTAGTGGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTTTGACGTTGGAGTCCA 3610  
Db |||||  
Qy 2563 GTTCAAGTAGTGGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTTTGACGTTGGAGTCCA 2622  
Db |||||  
Qy 3611 GGTCTTTTAATAGTGGACTCTGTTCCAAACCTGGAAACAACTCAACCTCTATCTCGGTCT 3670  
Db |||||  
Qy 2623 GGTCTTTTAATAGTGGACTCTGTTCCAAACCTGGAAACAACTCAACCTCTATCTCGGTCT 2682  
Db |||||  
Qy 3671 ATTCTTTTGAATTTAATAAGGGAATTTTCGCGATTTTCGCCCTTATGTTTAAATAAATGAGCTGA 3730  
Db |||||  
Qy 2683 ATTCTTTTGAATTTAATAAGGGAATTTTCGCGATTTTCGCCCTTATGTTTAAATAAATGAGCTGA 2742  
Db |||||  
Qy 3731 TTTAAACAAATATTTAAACCGGAATTTTAAACAAATATTTAAACGTTTAAACATTTTCGCCCTGATG 3790  
Db |||||  
Qy 2743 TTTAAACAAATATTTAAACCGGAATTTTAAACAAATATTTAAACGTTTAAACATTTT----- 2793  
Db |||||  
Qy 3791 CGGTATTTTCTCTTAAOCATCTGTGCGGTATTTTACACCGCATACAGGTGGCAGCTTTTC 3850  
Db |||||  
Qy 2794 -----AGGTGGCAGCTTTTC 2807  
Db |||||  
Qy 3851 GGGGAAATGTGCGGGAAACCCCTATTTGTTTATTTTCTAAATACATTTCAAAATATGTATC 3910  
Db |||||  
Qy 2808 GGGGAAATGTGCGGGAAACCCCTATTTGTTTATTTTCTAAATACATTTCAAAATATGTATC 2867  
Db |||||  
Qy 3911 CGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATAATTTGAAAGAAAGAGATATGA 3970  
Db |||||  
Qy 2868 CGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATAATTTGAAAGAAAGAGATATGA 2927  
Db |||||  
Qy 3971 GTATTTCAACATTTCCGTGTGCGCTTATTTCCCTTTTTCGCGCATTTTTCCTTTTTCCTTTT 4030  
Db |||||  
Qy 2928 GTATTTCAACATTTCCGTGTGCGCTTATTTCCCTTTTTCGCGCATTTTTCCTTTTTCCTTTT 2987  
Db |||||  
Qy 4031 TTGCTCACCAGAAACGCTGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGGACAGAG 4090  
Db |||||  
Qy 2988 TTGCTCACCAGAAACGCTGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGGACAGAG 3047  
Db |||||  
Qy 4091 TGGGTTACATCGAATCGAATCTCAACAGCGGTGAAGATCTTTGAGAGTTTTCGCCCGGAG 4150  
Db |||||  
Qy 3048 TGGGTTACATCGAATCGAATCTCAACAGCGGTGAAGATCTTTGAGAGTTTTCGCCCGGAG 3107  
Db |||||  
Qy 4151 AACGTTTTCATGATGAGCACTTTTAAAGTCTGCTATGTCGCGCGGTATTATCCCGTA 4210  
Db |||||  
Qy 3108 AACGTTTTCATGATGAGCACTTTTAAAGTCTGCTATGTCGCGCGGTATTATCCCGTA 3167  
Db |||||  
Qy 4211 TTGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTTCTCAGAAATGATCTGGTTG 4270  
Db |||||  
Qy 3168 TTGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTTCTCAGAAATGATCTGGTTG 3227  
Db |||||  
Qy 4271 AGTACTCACCAGTCAAGAAAGCATCTTAAGATGCGATGAGATGACAGTAAAGAAATTTATGA 4330  
Db |||||  
Qy 3228 AGTACTCACCAGTCAAGAAAGCATCTTAAGATGCGATGAGATGACAGTAAAGAAATTTATGA 3287  
Db |||||  
Qy 4331 GTGCTGCGCATTAACCATGAGTGAATCACTGCGCGCAACTTCTCTGACAAAGTGGAG 4390  
Db |||||  
Qy 3288 GTGCTGCGCATTAACCATGAGTGAATCACTGCGCGCAACTTCTCTGACAAAGTGGAG 3347  
Db |||||  
Qy 4391 GACCGAAGGAGCTAAACCGCTTTTTCGACAAACATGCGGGGATCATATGATCGCTTGTATC 4450  
Db |||||  
Qy 3348 GACCGAAGGAGCTAAACCGCTTTTTCGACAAACATGCGGGGATCATATGATCGCTTGTATC 3407  
Db |||||  
Qy 4451 GTTGGGAAACCGGAGCTGAATGAAGCATATCCAAAGAGAGGCTGACACCGAGATGCGCTG 4510  
Db |||||  
Qy 3408 GTTGGGAAACCGGAGCTGAATGAAGCATATCCAAAGAGAGGCTGACACCGAGATGCGCTG 3467  
Db |||||  
Qy 4511 TAGCAATGGGCAACCAAGTTGCGCAAACTTATTAACCTGGCGAACTACTTACTTAGCTTCC 4570  
Db |||||  
Qy 3468 TAGCAATGGGCAACCAAGTTGCGCAAACTTATTAACCTGGCGAACTACTTACTTAGCTTCC 3527  
Db |||||  
Qy 4571 GGCAACAAATTAATAGATGGAATGGAGCGGATAAAGTTTCAGGACCACTTTCTGCGCTCGG 4630  
Db |||||  
Qy 3528 GGCAACAAATTAATAGATGGAATGGAGCGGATAAAGTTTCAGGACCACTTTCTGCGCTCGG 3587  
Db |||||



```
QY 4631 CCCTTCGGCTGGCTGTTATTTGCTGATAAATCTGGAGCGGCTGAGCGTGGGTCTCGCG 4690
DB 3588 CCCTTCGGCTGGCTGTTATTTGCTGATAAATCTGGAGCGGCTGAGCGTGGGTCTCGCG 3647
QY 4691 GTATCATTCGACACTGGGCGCAGATGCTAAGCCCTCCCGCTATCGTAGTTATCTACACGA 4750
DB 3648 GTATCATTCGACACTGGGCGCAGATGCTAAGCCCTCCCGCTATCGTAGTTATCTACACGA 3707
QY 4751 CGGGAGTCAGGCAACTATGGATGAACGAATAAGACAGATCGCTGAGATAGGTGCCTCAC 4810
DB 3708 CGGGAGTCAGGCAACTATGGATGAACGAATAAGACAGATCGCTGAGATAGGTGCCTCAC 3767
QY 4811 TGATTAAGCACTGGTAAGCTCTCAGACCAAGTTTACTCATATATACCTTAGATGATTTAA 4870
DB 3768 TGATTAAGCACTGGTAAGCTCTCAGACCAAGTTTACTCATATATACCTTAGATGATTTAA 3827
QY 4871 AACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGATATCTCATGACCA 4930
DB 3828 AACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGATATCTCATGACCA 3887
QY 4931 AAATCCCTTAAAGTGTGTTTCCACTGAGCGTCAGACCCCGTAGAAAGATCAAAG 4990
DB 3888 AAATCCCTTAAAGTGTGTTTCCACTGAGCGTCAGACCCCGTAGAAAGATCAAAG 3947
QY 4991 GATCTCTTGAGATCCCTTTTTCGGCGGTAAATCTGCTGTGCAACCAAAAAACAC 5050
DB 3948 GATCTCTTGAGATCCCTTTTTCGGCGGTAAATCTGCTGTGCAACCAAAAAACAC 4007
QY 5051 CGCTACAGCGGTGGTTGTTGCCGATCAAGAGCTACCAACTCTTTTTCGGAAGTAA 5110
DB 4008 CGCTACAGCGGTGGTTGTTGCCGATCAAGAGCTACCAACTCTTTTTCGGAAGTAA 4067
QY 5111 CTGGCTTCAGCAGAGCGCAGATACCAAAATCTGCTTCTAGTGTAGCCGTAGTAGGCC 5170
DB 4068 CTGGCTTCAGCAGAGCGCAGATACCAAAATCTGCTTCTAGTGTAGCCGTAGTAGGCC 4127
QY 5171 ACCACTTCAGAACTCTGTAGCAGCGCTACATACCTCGCTCTGCTAACTCTGTACACG 5230
DB 4128 ACCACTTCAGAACTCTGTAGCAGCGCTACATACCTCGCTCTGCTAACTCTGTACACG 4187
QY 5231 TGCTCTGCTGCACTGCGATAGCTGCTTACCGGTTGGAATCAAGACGATAGTTAC 5290
DB 4188 TGCTCTGCTGCACTGCGATAGCTGCTTACCGGTTGGAATCAAGACGATAGTTAC 4247
QY 5291 CGGATTAAGCGCAGCGGTGCGGCTGAACCGGGGGTTGCTGCAACAGCCAGCTTGAGC 5350
DB 4248 CGGATTAAGCGCAGCGGTGCGGCTGAACCGGGGGTTGCTGCAACAGCCAGCTTGAGC 4307
QY 5351 GAACGACCTACACCGAACTGAGATACCTACAGCGTGAAGTATGAGAAAGCGCCACTTC 5410
DB 4308 GAACGACCTACACCGAACTGAGATACCTACAGCGTGAAGTATGAGAAAGCGCCACTTC 4367
QY 5411 CCGAAGGAGAAAGCGCAGGTATCCGTAAGCGGAGGTCGGAACAGGAGAGCGCA 5470
DB 4368 CCGAAGGAGAAAGCGCAGGTATCCGTAAGCGGAGGTCGGAACAGGAGAGCGCA 4427
QY 5471 CGAGGAGCTTCAGGGGGAACCGCTGTATCTTTATAGTCTGCTGCGGTTTCGCCACC 5530
DB 4428 CGAGGAGCTTCAGGGGGAACCGCTGTATCTTTATAGTCTGCTGCGGTTTCGCCACC 4487
QY 5531 TCTGACTTCAGCGTCCATTTTGTGATGCTGCTCAGGGGGGCGGAGCTATGGAAGAG 5590
DB 4488 TCTGACTTCAGCGTCCATTTTGTGATGCTGCTCAGGGGGGCGGAGCTATGGAAGAG 4547
QY 5591 CCAGCAACCGGCTTTTACGGTCTCTGCGCTTTTGTGCGCTTTTGTCTACATGTTCT 5650
DB 4548 CCAGCAACCGGCTTTTACGGTCTCTGCGCTTTTGTGCGCTTTTGTCTACATGTTCT 4607
QY 5651 TTCCTCGGTATCCCTCTGATTCGTGGAATAACCGTATTAACGCGCTTTGTAGTAGCTGATA 5710
DB 4608 TTCCTCGGTATCCCTCTGATTCGTGGAATAACCGTATTAACGCGCTTTGTAGTAGCTGATA 4667
QY 5711 CGCTCGCGCAGCCGACGACGAGCGCAGCGAGTCACTGAGCGAGGAGCGGAGAGC 5770
```

```
DB 4668 CCCTTCGGCTGGCTGTTATTTGCTGATAAATCTGGAGCGGCTGAGCGTGGGTCTCGCG 4727
QY 5771 GCCCAATACGCAAAACCGCTCTCCCGCGGTTTGGCCGATTCTATTAAATGCAG 5822
DB 4728 GCCCAATACGCAAAACCGCTCTCCCGCGGTTTGGCCGATTCTATTAAATGCAG 4779
```

## RESULT 7

```
US-09-991-209-33
; Sequence 33, Application US/09991209
; Publication No. US20030024009A1
; GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Langdon, Timothy
; APPLICANT: Morse, Phillip
; TITLE OF INVENTION: Manipulation of the Phenolic Acid
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes
; FILE REFERENCE: GC648-2
; CURRENT APPLICATION NUMBER: US/09/991,209
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/249,608
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 5034
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pJ05 vector
US-09-991-209-33
```

Query Match 47.5%; Score 2768; DB 10; Length 5034;

Best Local Similarity 96.8%; Pred. No. 0;

Matches 2895; Conservative 0; Mismatches 25; Indels 72; Gaps 4;

```
QY 2835 AAGCAGATCGTTCAAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCTCTTCCGGTC 2894
DB 1889 AAGCAGATCGTTCAAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCTCTTCCGGTC 1948
QY 2895 TTGCGATGATTATCATATATATTTCTGTTGAAATTTACGTTAAGCATGTAAATTAACATGT 2954
DB 1949 TTGCGATGATTATCATATATATTTCTGTTGAAATTTACGTTAAGCATGTAAATTAACATGT 2008
QY 2955 AATGCAATGAGTTATTTATGAGATGGTTTTTATGATTAGATGCTCCGCAATTAACATTT 3014
DB 2009 AATGCAATGAGTTATTTATGAGATGGTTTTTATGATTAGATGCTCCGCAATTAACATTT 2068
QY 3015 AATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTAATCCGCGCGGTGT 3074
DB 2069 AATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTAATCCGCGCGGTGT 2128
QY 3075 CATCTATGTTACTAGATTCGACCTCGAGGATGGGATCCGCGGCGCATGCGACGTGGGC 3134
DB 2129 CATCTATGTTACTAGATTCGATATTAC---AATTCACCTGGCGCTGCTTTTACACGTC 3191
QY 3135 CCAATTCGCGCTTATAGTAGTGTGATTAACCGCGCTCAGTCGCGCTGCTTTTACACGTC 2235
DB 2176 CCAATTCGCGCTTATAGTAGTGTGATTAACCGCGCTCAGTCGCGCTGCTTTTACACGTC 3251
QY 3192 GTGACTGGGAAACCCCTGGCGTTACCCAACTTAATCCCTTTCGAGCAGCATCCCTCTTCG 2295
DB 2236 GTGACTGGGAAACCCCTGGCGTTACCCAACTTAATCCCTTTCGAGCAGCATCCCTCTTCG 3311
QY 3252 CCAGCTGGCGTAAATAGCGAGAGGCGCGCACCGATCGCCCTTCCCAACAGTTCGCGAGCC 2355
DB 2296 CCAGCTGGCGTAAATAGCGAGAGGCGCGCACCGATCGCCCTTCCCAACAGTTCGCGAGCC 3370
QY 3312 TGAATGGCGCAAT--GGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGTGTGCTGCTTA 2415
DB 2356 TGAATGGCGCAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGTGTGCTGCTTA
```



Db 4521 TCTGACTTGAGCGTGGATTTTGTGATGCTTCGTGAGGGGGCGAGCGCTATGGAAGACG 4580  
Qy 5591 CCAGCAACGCGGCTTTTACGGTCTCTGSCCTTTTGTGCTGSCCTTTTGTCTCACATGTTCT 5650  
Db 4581 CCAGCAACGCGGCTTTTACGGTCTCTGSCCTTTTGTGCTGSCCTTTTGTCTCACATGTTCT 4640  
Qy 5651 TTCTGCGTTATCCCTGATCTCTGTGATTAACCGTATTACGCTTTGAGTGAGCTGATA 5710  
Db 4641 TTCTGCGTTATCCCTGATCTCTGTGATTAACCGTATTACGCTTTGAGTGAGCTGATA 4700  
Qy 5711 CCGCTCGCCGAGCGCAACCGAGCGAGCGAGTCAGTGAAGTGAAGCGGAAGGC 5770  
Db 4701 CCGCTCGCCGAGCGCAACCGAGCGAGCGAGTCAAGTGAAGTGAAGCGGAAGGC 4760  
Qy 5771 GCCCAATACGCAACCGCCCTCTCCCGCGCGTTGGCCGATTCATTAATGCGAG 5822  
Db 4761 GCCCAATACGCAACCGCCCTCTCCCGCGCGTTGGCCGATTCATTAATGCGAG 4812

RESULT 8  
US-09-991-209-36  
; Sequence 36, Application US/09991209  
; Publication No. US20030024009A1  
; GENERAL INFORMATION:  
; APPLICANT: Dunn-Coleman, Nigel  
; APPLICANT: Langdon, Timothy  
; APPLICANT: Morse, Phillip  
; TITLE OF INVENTION: Manipulation of the Phenolic Acid  
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted  
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes  
; FILE REFERENCE: GC648-2  
; CURRENT APPLICATION NUMBER: US/09/991,209  
; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: US 60/249,608  
; PRIOR FILING DATE: 2000-11-17  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 36  
; LENGTH: 5164  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pPQ10.1 vector  
US-09-991-209-36

Query Match 47.5%; Score 2768; DB 10; Length 5164;  
Best Local Similarity 96.8%; Pred. No. 0;  
Matches 2895; Conservative 0; Mismatches 25; Indels 72; Gaps 4;

Qy 2835 AAGCAGATCGTTCAACATTTGGCAATTAAGTTTCTTAAGATTGAATTCCTGTTCCGGTC 2894  
Db 778 AAGCAGATCGTTCAACATTTGGCAATTAAGTTTCTTAAGATTGAATTCCTGTTCCGGTC 837  
Qy 2895 TTGGGATGATTATCATATAATTTCTGTTGAATTAAGTTTGAAGTGAATTAATTAACATGT 2954  
Db 838 TTGGGATGATTATCATATAATTTCTGTTGAATTAAGTTTGAAGTGAATTAATTAACATGT 897  
Qy 2955 AATGCATGACGTATTTATGAGATGGGTTTATGATTAAGTTCCTGCAATTAATTAATTT 3014  
Db 898 AATGCATGACGTATTTATGAGATGGGTTTATGATTAAGTTCCTGCAATTAATTAATTT 957  
Qy 3015 AATACGGATAGAAAACAAATATAGCGCGCAAACTAGGATAAATTAATGCGCGGGTGT 3074  
Db 958 AATACGGATAGAAAACAAATATAGCGCGCAAACTAGGATAAATTAATGCGCGGGTGT 1017  
Qy 3075 CATCTATGTTACTAGATCGACCTCGAGCGATGGGATCCGCGCGCATGCGACGTGCGGC 3134  
Db 1018 CATCTATGTTACTAGATCGATTAAGCTTCTAGAGCGCGCGGTGGAGC-----T 1064  
Qy 3135 CCAATTCGCCCTATAGTGAAGTGAATTAAC-----AATTCATCGCGCGGTGTTTAAACAGTC 3191  
Db 1065 CCAATTCGCCCTATAGTGAAGTGAATTAACCGCGCTCACTGCGCGCTGTTTAAACAGTC 1124

Qy 3192 GTGACTGGGAAAACCCCTGGCGTTACCCAACTTAATTCGCCCTTGAGAGCAATCCCTCTTCG 3251  
Db 1125 GTGACTGGGAAAACCCCTGGCGTTACCCAACTTAATTCGCCCTTGAGAGCAATCCCTCTTCG 1184  
Qy 3252 CCAGCTGGCGTAAATAGCGAAGAGCGCCGACCGATCGCCCTTCCCAACAGTTGGCGAGCC 3311  
Db 1185 CCAGCTGGCGTAAATAGCGAAGAGCGCCGACCGATCGCCCTTCCCAACAGTTGGCGAGCC 1244  
Qy 3312 TGAATGCGCGAAT- GGAACGCGCCCTGTAGCGCGCATTAAGCGCGCGGGTGTGTGTTA 3370  
Db 1245 TGAATGCGCGAATGGAACGCGCCCTGTAGCGCGCATTAAGCGCGCGGGTGTGTGTTA 1304  
Qy 3371 CGCGAGCGTGACGCGCTACACTTGCAGCGCCCTAGCGCGCGCTCTTTCGCTTCTTC 3430  
Db 1305 CGCGAGCGTGACGCGCTACACTTGCAGCGCCCTAGCGCGCGCTCTTTCGCTTCTTC 1364  
Qy 3431 CTTTCTTTCTCGGCAACGTTTCGCGGCTTTCGCGTCAAGCTCTAAATCGGCGGCTCCCTT 3490  
Db 1365 CTTTCTTTCTCGGCAACGTTTCGCGGCTTTCGCGTCAAGCTCTAAATCGGCGGCTCCCTT 1424  
Qy 3491 TAGGGTTCGATTTAGAGCTTTACGCACTCGACCGCAAAAAAATCTGATTTGGGTGATG 3550  
Db 1425 TAGGGTTCGATTTAGAGCTTTACGCACTCGACCGCAAAAAAATCTGATTTGGGTGATG 1484  
Qy 3551 GTTCACTAGTGGGCCATCGCCCTGATAGAGCGTTTTCGCGCCCTTTGAGCTGAGTCCA 3610  
Db 1485 GTTCACTAGTGGGCCATCGCCCTGATAGAGCGTTTTCGCGCCCTTTGAGCTGAGTCCA 1544  
Qy 3611 CGTTCTTTAATAGTGAAGTCTGTTTCCAAACTGGAACAACACTCAACCTATCTCGGTCT 3670  
Db 1545 CGTTCTTTAATAGTGAAGTCTGTTTCCAAACTGGAACAACACTCAACCTATCTCGGTCT 1604  
Qy 3671 ATTCTTTGATTTAAGGATTTTTCGCGATTTTCGCGCTATTTGTTTAAAAAATGAGCTGA 3730  
Db 1605 ATTCTTTGATTTAAGGATTTTTCGCGATTTTTCGCGCTATTTGTTTAAAAAATGAGCTGA 1664  
Qy 3731 TTTAAACAATATTTAAGCGCAATTTTAAACAATAATTAAGCTTTTACAAATTTTCGCTGATG 3790  
Db 1665 TTTAAACAATATTTAAGCGCAATTTTAAACAATAATTAAGCTTTTACAAATTTTCGCTGATG 1715  
Qy 3791 CGGTATTTTCTCTTACGCACTGTGCGGTATTTTACACCGCATACAGGTGGCACTTTTC 3850  
Db 1716 -----AGGTGGCACTTTTC 1729  
Qy 3851 GGGGAAATGTGCGGGAACCCCTATTTGTTTATTTTCTAAATACATTTCAAAATGATATC 3910  
Db 1730 GGGGAAATGTGCGGGAACCCCTATTTGTTTATTTTCTAAATACATTTCAAAATGATATC 1789  
Qy 3911 CGCTCATGACACAATACCTGATAAATGCTTCAATAATTTGAAAGGAGGAGATGA 3970  
Db 1790 CGCTCATGACACAATACCTGATAAATGCTTCAATAATTTGAAAGGAGGAGATGA 1849  
Qy 3971 GTATTCAACATTTTCGCTGCGCCCTTATTCCTTTTTTTCGCGCATTTTTCCTCTGTTT 4030  
Db 1850 GTATTCAACATTTTCGCTGCGCCCTTATTCCTTTTTTTCGCGCATTTTTCCTCTGTTT 1909  
Qy 4031 TTGCTCAACCCAGAAAACCTGCTGAAGTAAAGATGCTGAAGATCAGTTGGGTGCAACGAG 4090  
Db 1910 TTGCTCAACCCAGAAAACCTGCTGAAGTAAAGATGCTGAAGATCAGTTGGGTGCAACGAG 1969  
Qy 4091 TGGGTTACATCGAACTCGATCTCAACAGCGGTAGATCTTTGAGAGTTTTCGCGCCGAG 4150  
Db 1970 TGGGTTACATCGAACTCGATCTCAACAGCGGTAGATCTTTGAGAGTTTTCGCGCCGAG 2029  
Qy 4151 AACGTTTTTCAATGATGAGCACTTTTAAAGTTTCTGCTATGTTGGCGCGGTATTTATCCCGTA 4210  
Db 2030 AACGTTTTTCAATGATGAGCACTTTTAAAGTTTCTGCTATGTTGGCGCGGTATTTATCCCGTA 2089  
Qy 4211 TTGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTTCTCAGATGACTTGGTTG 4270  
Db 2090 TTGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTTCTCAGATGACTTGGTTG 2149

4271 AGTACTCACCAGTCACAGAAAGCACTTACGGATGCGATGACAGTAAGAGAAATTATGCA 4330  
Db AGTACTCACCAGTCACAGAAAGCACTTACGGATGCGATGACAGTAAGAGAAATTATGCA 2209  
4331 GTGCTGCATAAACCATGAGTGATAAACAATGCGGCCAACTTACTTCTGACAAACGATCGGAG 4390  
Db GTGCTGCATAAACCATGAGTGATAAACAATGCGGCCAACTTACTTCTGACAAACGATCGGAG 2269  
4391 GACCGAAGGAGCTAACCGCTTTTGTGACAAACATGCGGGGATCATGTAACTCGCCTTGATC 4450  
Db GACCGAAGGAGCTAACCGCTTTTGTGACAAACATGCGGGGATCATGTAACTCGCCTTGATC 2329  
4451 GTTGGGACCGGAGCTGAATGAGCCATACCAACGAGCGCGTGCACACGATGCGTG 4510  
Db GTTGGGACCGGAGCTGAATGAGCCATACCAACGAGCGCGTGCACACGATGCGTG 2389  
4511 TAGCAATGGCAACAAACGTTGCGCAAACTATTAACCTGGCGAACTTACTTCTAGCTTCCC 4570  
Db TAGCAATGGCAACAAACGTTGCGCAAACTATTAACCTGGCGAACTTACTTCTAGCTTCCC 4570  
2390 TAGCAATGGCAACAAACGTTGCGCAAACTATTAACCTGGCGAACTTACTTCTAGCTTCCC 2449  
4571 GGCAACAAATTAATAGACTGGAAGGCGGATAAAGTTGACGACCACTTCTGCGCTCGG 4630  
Db GGCAACAAATTAATAGACTGGAAGGCGGATAAAGTTGACGACCACTTCTGCGCTCGG 2509  
4631 CCTTCCGGCTGGCTGTTTATGCTGATAAATCTGGAGCGGAGCGGTGCTCGCG 4690  
Db CCTTCCGGCTGGCTGTTTATGCTGATAAATCTGGAGCGGAGCGGTGCTCGCG 2569  
4691 GTATCATTCGAGCACTGGGCGCAGATGTAAGCCCTCCCGTATCGTAGTTATCTACACGA 4750  
Db GTATCATTCGAGCACTGGGCGCAGATGTAAGCCCTCCCGTATCGTAGTTATCTACACGA 2629  
4751 CGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCTTCAC 4810  
Db CGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCTTCAC 2689  
4811 TGATTAAGCAATGGTAATCTGTCAGACCAAGTTTACTCATATATCTTACTGATGATTTAA 4870  
Db TGATTAAGCAATGGTAATCTGTCAGACCAAGTTTACTCATATATCTTACTGATGATTTAA 2749  
4871 AACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTAAATCTCATGACCA 4930  
Db AACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTAAATCTCATGACCA 2809  
4931 AAATCCCTTAACTGAGTCTTCTGCTCACTGAGCGTCAGACCCCGTAGAAGATCAAG 4990  
Db AAATCCCTTAACTGAGTCTTCTGCTCACTGAGCGTCAGACCCCGTAGAAGATCAAG 2869  
4991 GATCTCTTGAGATCCTTTTCTGCGGTAATCTGCTGTCGAAACAAAACCCAC 5050  
Db GATCTCTTGAGATCCTTTTCTGCGGTAATCTGCTGTCGAAACAAAACCCAC 2929  
5051 CGCTACAGCGGTGTTTGTGCGGATCAAGAGCTACCAACTCTTTTCCGAAGTAA 5110  
Db CGCTACAGCGGTGTTTGTGCGGATCAAGAGCTACCAACTCTTTTCCGAAGTAA 2989  
5111 CTGCTCTCAGCAGCGCAGATACCAATCTGCTTCTAGTGTAGCGGTAGTTAGGCC 5170  
Db CTGCTCTCAGCAGCGCAGATACCAATCTGCTTCTAGTGTAGCGGTAGTTAGGCC 3049  
5171 ACCACTTCAAGAACTCTGTAGACCGCTCATACCTCGCTCTGCTGATCTGTTTACAG 5230  
Db ACCACTTCAAGAACTCTGTAGACCGCTCATACCTCGCTCTGCTGATCTGTTTACAG 3109  
5231 TGCTGCTGCGGATGAGTCTGCTTACCGGGTTGGACTCAAGACGATAGTTAC 5290  
Db TGCTGCTGCGGATGAGTCTGCTTACCGGGTTGGACTCAAGACGATAGTTAC 3169  
5291 CGGATAGGCGCAGCGGTGCGGCTGAACCGGGGTTGCTGACACAGCCCGCTTGGAGC 5350  
Db CGGATAGGCGCAGCGGTGCGGCTGAACCGGGGTTGCTGACACAGCCCGCTTGGAGC 3229  
5351 GAAACGACCTACACCGAACTGAGATACCTACAGCGGTGAGCTATGAGAAAGCGCCACGCTTC 5410

3230 GAACGACCTTACCGAACTGAGATACCTTACAGCTGAGCTATGAGAAAGCGCCACGCTTC 3289  
Qy CCGAAGGAGAAAGCGCGACAGAGTATCCGGTAAAGCGGAGGCTCGGAAACAGGAGCGCA 5470  
Db CCGAAGGAGAAAGCGCGACAGAGTATCCGGTAAAGCGGAGGCTCGGAAACAGGAGCGCA 3349  
5471 CGAGGAGCTTCCAGGGGGGAAACCGCTGCTATCTTTATGATCTGCTCGGTTTCCGCAAC 5530  
Db CGAGGAGCTTCCAGGGGGGAAACCGCTGCTATCTTTATGATCTGCTCGGTTTCCGCAAC 3409  
5531 TCTGACTTGAAGCGTGCATTTTGTGATGCTGCTCAGGGGGGCGGAGCCTATGGAACAG 5590  
Db TCTGACTTGAAGCGTGCATTTTGTGATGCTGCTCAGGGGGGCGGAGCCTATGGAACAG 3469  
5591 CCGAACAGCGGCGCTTTTACGGTTCCTGGCTTTTGTGCTGCTTCTGCTCAGATGTTCT 5650  
Db CCGAACAGCGGCGCTTTTACGGTTCCTGGCTTTTGTGCTGCTTCTGCTCAGATGTTCT 3529  
5651 TTCTGCTGTTATCCCTGATCTGCTGATTAACCGTATTTACCGCTTTTGTGAGTGAAGTATA 5710  
Db TTCTGCTGTTATCCCTGATCTGCTGATTAACCGTATTTACCGCTTTTGTGAGTGAAGTATA 3589  
5711 CCGCTGCGCGACCGCAACCGACCGAGCGCAGAGTCAAGTGAAGCGGAGAGC 5770  
Db CCGCTGCGCGACCGCAACCGACCGAGCGCAGAGTCAAGTGAAGCGGAGAGC 3649  
5771 GCCCAATACGCAAAACCGCTCTCCCGCGCTTCCCGCGCTTCCCGCGCTTCAATTAATGAG 5822  
Db GCCCAATACGCAAAACCGCTCTCCCGCGCTTCCCGCGCTTCAATTAATGAG 3701

RESULT 9  
US-09-991-209-25  
; Sequence 25, Application US/09991209  
; Publication No. US20030024009A1  
; GENERAL INFORMATION:  
; APPLICANT: Dunn-Coleman, Nigel  
; APPLICANT: Langdon, Timothy  
; APPLICANT: Morse, Phillip  
; TITLE OF INVENTION: Manipulation of the Phenolic Acid  
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted  
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes  
; FILE REFERENCE: GC648-2  
; CURRENT APPLICATION NUMBER: US/09/991,209  
; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: US 60/249,608  
; PRIOR FILING DATE: 2000-11-17  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25  
; LENGTH: 5277  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pTP5-1 vector  
US-09-991-209-25

Query Match 47.5% Score 2768; DB 10; Length 5277;  
Best Local Similarity 96.8%; Pred. No. 0;  
Matches 2895; Conservative 0; Mismatches 25; Indels 72; Gaps 4;  
Qy 2835 AAGCAGATCGTTCAAAACATTTGGCAATTAAGTTTCTTAAGATTCGATCTGTCGGGTC 2894  
Db 894 AAGCAGATCGTTCAAAACATTTGGCAATTAAGTTTCTTAAGATTCGATCTGTCGGGTC 953  
Qy 2895 TTGGATGATTCATTAATTTCTGTTGAATTCGTTAAGCATGTAATTAATTAACATGT 2954  
Db 954 TTGGATGATTCATTAATTTCTGTTGAATTCGTTAAGCATGTAATTAATTAACATGT 1013  
Qy 2955 AATGATGATGTTTATTTATGAGATGTTTATGATTTAGAGTCCCGCAATATATACATTT 3014  
Db 1014 AATGATGATGTTTATTTATGAGATGTTTATGATTTAGAGTCCCGCAATATATACATTT 1073

3015 AATPACGGATAGAAAAAATAATAGCGCGCAAACTAGGATAAAATTATCGCGCGGTGT 3074  
1074 AATACGGATAGAAAAAATAATAGCGCGCAAACTAGGATAAAATTATCGCGCGGTGT 1133  
3075 CATCTATGTTACTAGATCGACCTGCGAGGATGGGATCGCGCGCGATGCGAGCTCGGGC 3134  
1134 CATCTATGTTACTAGATCGGATAAGCTTTCTAGAGCGCGCGGTGGAGC-----T 1180  
3135 CCAATTCGCCCTATAGTAGAGTGGTATTAC---AATTCACTGGCGCGTGGTTTATCAACGTC 3191  
1181 CCAATTCGCCCTATAGTAGAGTGGTATTACGGCGGCTCACTGGCGGTGGTTTATCAACGTC 1240  
3192 GTGACTGGGAAAAACCTTGGCGTTTACCCAACTTAATCGCTTTCAGGACATTCGCCCTTTCG 3251  
1241 GTGACTGGGAAAAACCTTGGCGTTTACCCAACTTAATCGCTTTCAGGACATTCGCCCTTTCG 1300  
3252 CAGCTGGCGTAAATAGCGAGAGGCGCGACCGATCGCCCTCCACACAGTTCGGCAGCC 3311  
1301 CAGCTGGCGTAAATAGCGAGAGGCGCGACCGATCGCCCTCCACACAGTTCGGCAGCC 1360  
3312 TGAATGGCGAAT--GGACGCGCCCTGTAGCGCGCGCATTAAGCGCGCGGTGGTGGTTA 3370  
1361 TGAATGGCGAATGGGACGCGCCCTGTAGCGCGCGCATTAAGCGCGCGGTGGTGGTTA 1420  
3371 CGCGAGCGTGAACGCTTACACTTTGCCAGCGCCCTAGCGCGCGCTCTCTTTCGCTTCCTCC 3430  
1421 CGCGAGCGTGAACGCTTACACTTTGCCAGCGCCCTAGCGCGCGCTCTCTTTCGCTTCCTCC 1480  
3431 CTTCTCTTCTCGCAGCTTTCGCGGCTTTCGCGTCAAGCTCTAATCGGGGCTCCCTT 3490  
1481 CTTCTCTTCTCGCAGCTTTCGCGGCTTTCGCGTCAAGCTCTAATCGGGGCTCCCTT 1540  
3491 TAGGGTTCCGATTTAGAGCTTTTACGCGCACTTCGACCGCGCAAAACTTGATTGGGTGATG 3550  
1541 TAGGGTTCCGATTTAGTGTCTTACGCGCACTTCGACCGCGCAAAACTTGATTGGGTGATG 1600  
3551 GTTCAAGTGTGGCCCATCGCCCTGAAGAGCGTTTTCGCGCTTTCGCGCTTTCGCGCTTTCG 3610  
1601 GTTCAAGTGTGGCCCATCGCCCTGAAGAGCGTTTTCGCGCTTTCGCGCTTTCGCGCTTTCG 1660  
3611 CGTTCTTTAATAGTGGACTTTCGTTCCAACTGGAACTGGAACCACTCAACCTCATCTCGGCT 3670  
1661 CGTTCTTTAATAGTGGACTTTCGTTCCAACTGGAACTGGAACCACTCAACCTCATCTCGGCT 1720  
3671 ATTCTTTTGAATTAAGGGAATTTTCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCT 3730  
1721 ATTCTTTTGAATTAAGGGAATTTTCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCT 1780  
3731 TTTAAACAAATATTTAAACGGAATTTTAAACAAATATTTAAACGGAATTTTAAACGGAATTT 3790  
1781 TTTAAACAAATATTTAAACGGAATTTTAAACAAATATTTAAACGGAATTTTAAACGGAATTT 1831  
3791 CGGTATTTCTCTTACGCTCTGCGGTATTTTCAACCGCATACAGGTGGCGACTTTTTC 3850  
1832 -----AGGTGGCGACTTTTC 1845  
3851 GGCGAATGTGCGGGAACCCCTATTGTTTATTTTCTAAATACATTTCAAATATGATC 3910  
1846 GGCGAATGTGCGGGAACCCCTATTGTTTATTTTCTAAATACATTTCAAATATGATC 1905  
3911 CGCTCATGACAAATACCCCTGATTAATGCTTCAATTAATTTGAAAGGAAGGATGTA 3970  
1906 CGCTCATGACAAATACCCCTGATTAATGCTTCAATTAATTTGAAAGGAAGGATGTA 1965  
3971 GTATTCAACATTTTCGCTGCGCCCTTATCCCTTTTTCGCGCATTTTTCCTTCTGTTT 4030  
1966 GTATTCAACATTTTCGCTGCGCCCTTATCCCTTTTTCGCGCATTTTTCCTTCTGTTT 2025  
4031 TTGCTCAACGAGAAACGCTGTTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCGACGAG 4090  
2026 TTGCTCAACGAGAAACGCTGTTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCGACGAG 2085

4091 TGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTCTTGAGAGTGTTCGCGCCCGAAG 4150  
2086 TGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTCTTGAGAGTGTTCGCGCCCGAAG 2145  
4151 AAGTTTTTCCAATGATGAGCACATTTTAAAGTTCTGCTATCTGCGCGGTAATTTATCCCGTA 4210  
2146 AAGTTTTTCCAATGATGAGCACATTTTAAAGTTCTGCTATCTGCGCGGTAATTTATCCCGTA 2205  
4211 TTGACGCCGGGCAAGAGCAACTCGGTCGCGCATACACTATTCTCAGAAATGACTTGGTTG 4270  
2206 TTGACGCCGGGCAAGAGCAACTCGGTCGCGCATACACTATTCTCAGAAATGACTTGGTTG 2265  
4271 AGTACTCACAGTACAGAAAAGCATCTTACGATGCGATGACAGTAAGAAATTAATGCA 4330  
2266 AGTACTCACAGTACAGAAAAGCATCTTACGATGCGATGACAGTAAGAAATTAATGCA 2325  
4331 GTGCTGCCATAACCATGAGTAAACACTGCGCGCAACTTCTCTGCAACAGATCCGAG 4390  
2326 GTGCTGCCATAACCATGAGTAAACACTGCGCGCAACTTCTCTGCAACAGATCCGAG 2385  
4391 GACCGAAGGAGCTAACCGCTTTTTCACAAATGCGGATCATGTAACCTGCGCTTGATC 4450  
2386 GACCGAAGGAGCTAACCGCTTTTTCACAAATGCGGATCATGTAACCTGCGCTTGATC 2445  
4451 GTTGGGAAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGAACCAACGATGCTG 4510  
2446 GTTGGGAAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGAACCAACGATGCTG 2505  
4511 TAGCAATGCGCAACAACTGTTGCGCAAACTATTAACTGCGCAAACTTCTCTAGCTTCCC 4570  
2506 TAGCAATGCGCAACAACTGTTGCGCAAACTATTAACTGCGCAAACTTCTCTAGCTTCCC 2565  
4571 GGCAACAAATTAATAGACTGCGAGCGGATAAAGTTGCGAGCACTTCTGCGCTCGG 4630  
2566 GGCAACAAATTAATAGACTGCGAGCGGATAAAGTTGCGAGCACTTCTGCGCTCGG 2625  
4631 CCCTTCGCGCTGCGTGGTTTATTTGCTGATTAATTTCTGAGCGCGGTGAGGTTCTCGCG 4690  
2626 CCCTTCGCGCTGCGTGGTTTATTTGCTGATTAATTTCTGAGCGCGGTGAGGTTCTCGCG 2685  
4691 GTATCATTTGAGCACTTGGGCGCAGATGTAAGCCCTCCCGTATCGTAGTTATCTACACGA 4750  
2686 GTATCATTTGAGCACTTGGGCGCAGATGTAAGCCCTCCCGTATCGTAGTTATCTACACGA 2745  
4751 CGGGAGTCAAGCAACTATGATGATGAACGAAATAGACAGATCGCTGAGATAGGTCCTCAC 4810  
2746 CGGGAGTCAAGCAACTATGATGATGAACGAAATAGACAGATCGCTGAGATAGGTCCTCAC 2805  
4811 TGAATTAAGCAATTTGATTAACCTGTCAGACCAAGTTTACTCATATATCTTTAGATTGATTTAA 4870  
2806 TGAATTAAGCAATTTGATTAACCTGTCAGACCAAGTTTACTCATATATCTTTAGATTGATTTAA 2865  
4871 AACTTCATTTTAAATTTAAAGGATCTAGTGAAGATCTTTTGTGATAATCTCATGACCA 4930  
2866 AACTTCATTTTAAATTTAAAGGATCTAGTGAAGATCTTTTGTGATAATCTCATGACCA 2925  
4931 AATTCCTTAAACGTGAGTTTTCGTCAGCTGAGCGTCAGACCCGCTAGAAAGATCAAG 4990  
2926 AATTCCTTAAACGTGAGTTTTCGTCAGCTGAGCGTCAGACCCGCTAGAAAGATCAAG 2985  
4991 GATCTTCTTGAGATCTTTTTCGCGGTAATCTGCTGCTTTCGCAACAAACAAACCAAC 5050  
2986 GATCTTCTTGAGATCTTTTTCGCGGTAATCTGCTGCTTTCGCAACAAACAAACCAAC 3045  
5051 CGCTACAGCGGTGGTTGTTTCGCGGATCAAGAGCTACCAACTCTTTTTCGAGGATA 5110  
3046 CGCTACAGCGGTGGTTGTTTCGCGGATCAAGAGCTACCAACTCTTTTTCGAGGATA 3105  
5111 CTGGCTTACAGAGCGGAGATACCAAAATCTGCTCTTCTAGTGTAGCGGTAGTGGCC 5170  
3106 CTGGCTTACAGAGCGGAGATACCAAAATCTGCTCTTCTAGTGTAGCGGTAGTGGCC 3165  
5171 ACCACTTCAAGAACTCTGTAGCACCGCTACATACCTGCTCTGCTAATCTCTGTACCG 5230

Db 3166 ACCACTTCAGAACTCTCTGTAGCACCGCTACATACCTCGCTCTGCTATCTCTGTACCG 3225  
Qy 5231 TGGCTGTGCGCAGTGGCGATAGTCTGTCTTACCGGGTTGGACTCAAGACGATAGTTAC 5290  
Db 3226 TGGCTGTGCGCAGTGGCGATAGTCTGTCTTACCGGGTTGGACTCAAGACGATAGTTAC 3285  
Qy 5291 CGGATAGCGCGAGCGGTGGCGTGAACGGGGGTTCTGTCACACAGCCAGCTTGGAGC 5350  
Db 3286 CGGATAGCGCGAGCGGTGGCGTGAACGGGGGTTCTGTCACACAGCCAGCTTGGAGC 3345  
Qy 5351 GAAAGGAGTACACCGGAACTGAGATACCTACAGCTGAGCTATGAGAAAGCGCCAGCTTC 5410  
Db 3346 GAAAGGAGTACACCGGAACTGAGATACCTACAGCTGAGCTATGAGAAAGCGCCAGCTTC 3405  
Qy 5411 CCGAAGGAGAAAGCGGAGGATATCCGGTAAGCGGAGGGTTCGGAACAGSAGAGCGCA 5470  
Db 3406 CCGAAGGAGAAAGCGGAGGATATCCGGTAAGCGGAGGGTTCGGAACAGSAGAGCGCA 3465  
Qy 5471 CGAGGAGCTTCAGGGGGAAGCGCTGGTATCTTTATAGTCTCTGCGGGTTTCGCCACC 5530  
Db 3466 CGAGGAGCTTCAGGGGGAAGCGCTGGTATCTTTATAGTCTCTGCGGGTTTCGCCACC 3525  
Qy 5531 TCTGACTTGAGCGTCTGATTTTGTGATGCTCTGACGGGGGAGCGCTATGGAAGAACG 5590  
Db 3526 TCTGACTTGAGCGTCTGATTTTGTGATGCTCTGACGGGGGAGCGCTATGGAAGAACG 3585  
Qy 5591 CCAGCAACCGCGGCTTTTACGGTTTCTGCGCTTTTGTGCTGAGCTATGGAAGAACG 5650  
Db 3586 CCAGCAACCGCGGCTTTTACGGTTTCTGCGCTTTTGTGCTGAGCTATGGAAGAACG 3645  
Qy 5651 TTCCTGCTTATCCCTGATCTGTGGATAACCGTATTTACCGCTTTGAGTGGCTGATA 5710  
Db 3646 TTCCTGCTTATCCCTGATCTGTGGATAACCGTATTTACCGCTTTGAGTGGCTGATA 3705  
Qy 5711 CGCTGCGCGAGCGGACCGAGCGAGCTAGTCAAGTGGAGCGGAGCGGAGCGGAGGAGC 5770  
Db 3706 CGCTGCGCGAGCGGACCGAGCGAGCTAGTCAAGTGGAGCGGAGCGGAGCGGAGGAGC 3765  
Qy 5771 GCCCAATACGCAACCGCTCTCCCCGCGGTGGCCGATTCATTAATGCAG 5822  
Db 3766 GCCCAATACGCAACCGCTCTCCCCGCGGTGGCCGATTCATTAATGCAG 3817

RESULT 10  
US-09-991-209-38  
; Sequence 38, Application US/09991209  
; Publication No. US20030024009A1  
; GENERAL INFORMATION:  
; APPLICANT: Dunn-Coleman, Nigel  
; APPLICANT: Langdon, Timothy  
; APPLICANT: Morse, Philip  
; TITLE OF INVENTION: Manipulation of the Phenolic Acid  
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted  
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes  
; FILE REFERENCE: GC648-2  
; CURRENT APPLICATION NUMBER: US/09/991,209  
; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: US 60/249,608  
; PRIOR FILING DATE: 2000-11-17  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 38  
; LENGTH: 5295  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pUG4 vector  
US-09-991-209-38  
Query Match .47.5%; Score 2768; DB 10; Length 5295;  
Best Local Similarity 96.8%; Pred No. 0;  
Matches 2895; Conservative 0; Mismatches 25; Indels 72; Gaps 4;

Qy 2835 AAGCAGATCGTTCAAAACATTTGGCAATTAAGTTTCTTAAGATTGAATCCTGTTGCCGGTC 2894  
Db 912 AAGCAGATCGTTCAAAACATTTGGCAATTAAGTTTCTTAAGATTGAATCCTGTTGCCGGTC 971  
Qy 2895 TTGCGATGATTAATCATATAAATTTCTGTTGAAATTAAGTTTAAGCATGTAAATAAATAACATGT 2954  
Db 972 TTGCGATGATTAATCATATAAATTTCTGTTGAAATTAAGTTTAAGCATGTAAATAAATAACATGT 1031  
Qy 2955 AATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATATATACATTT 3014  
Db 1032 AATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATATATACATTT 1091  
Qy 3015 AATACCGATAGAAAACAAAATATAGCGGCAAACTAGGATAAATATATCCGGCGGGTGT 3074  
Db 1092 AATACCGATAGAAAACAAAATATAGCGGCAAACTAGGATAAATATATCCGGCGGGTGT 1151  
Qy 3075 CATCTATGTTACTAGATCGACCTCGAGGATCCGGCCGCGCATCGCAGCTCGGGC 3134  
Db 1152 CATCTATGTTACTAGATCGATGAAGCTTCTAGAGCGGCGGGTGAGC-----T 1198  
Qy 3135 CCAATTCGCCCTATATAGTGAAGTGTATTAC---AATTCATCGCCCGTCTGTTTACAAAGTC 3191  
Db 1199 CCAATTCGCCCTATATAGTGAAGTGTATTACGGCGGCTCACTGCGCGTCTGTTTACAAAGTC 1258  
Qy 3192 GTGACTGGGAAAACCCCTGGCGTTACCCAACTTAATGCGCTTCGAGCAATCCCTTTCG 3251  
Db 1259 GTGACTGGGAAAACCCCTGGCGTTACCCAACTTAATGCGCTTCGAGCAATCCCTTTCG 1318  
Qy 3252 CCAGCTGCGGTAATAGCGAAGAGGCGCGCACGATCGCCCTTCCAAAGTGTGCGCAGCC 3311  
Db 1319 CCAGCTGCGGTAATAGCGAAGAGGCGCGCACGATCGCCCTTCCAAAGTGTGCGCAGCC 1378  
Qy 3312 TGAATGGCGAAT-GGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGTGTGGTTA 3370  
Db 1379 TGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGTGTGGTTA 1438  
Qy 3371 CGCGCAGGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCTTTCGCTTCTTCC 3430  
Db 1439 CGCGCAGGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCTTTCGCTTCTTCC 1498  
Qy 3431 CTTCTCTTCTGCGCACGTTTCCGCGCTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTT 3490  
Db 1499 CTTCTCTTCTGCGCACGTTTCCGCGCTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTT 1558  
Qy 3491 TAGGGTCCGATTTTAGAGCTTTACGCGACCTCGACCGCAAAAACCTTGAATTTGGGTGATG 3550  
Db 1559 TAGGGTCCGATTTTAGTGTCTTTACGCGACCTCGACCGCAAAAACCTTGAATTTAGGTTGATG 1618  
Qy 3551 GTTCAAGTGTGGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTTGACGTTGGAGTCCA 3610  
Db 1619 GTTCAAGTGTGGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTTTGAAGTTGGAGTCCA 1678  
Qy 3611 CGTCTCTTAATAGTGGACTCTTGTTCAAACTGGAACACACTCAACCTCTATCTCGGTCT 3670  
Db 1679 CGTCTCTTAATAGTGGACTCTTGTTCAAACTGGAACACACTCAACCTCTATCTCGGTCT 1738  
Qy 3671 ATTCTTTGATTTAAGGATTTTTCGCGATTTTCGCGCTTATGTTTAAATAATAGACTGA 3730  
Db 1739 ATTCTTTGATTTAAGGATTTTTCGCGATTTTCGCGCTTATGTTTAAATAATAGACTGA 1798  
Qy 3731 TTTACAAATATTTAAACGGAATTTTACAAATATTAAGGTTTACAAATTTCCGCTGATG 3790  
Db 1799 TTTACAAATATTTAAACGGAATTTTACAAATATTAAGGTTTACAAATTTTACAAATTT 1849  
Qy 3791 CGGTATTTTCTCTTACGATCTGTGCGGTATTTTACACCGCATACAGGTGGCCTTTTC 3850  
Db 1850 -----AGGTGGCAGCTTTTC 1863  
Qy 3851 GGGGAAATGTGCGGGAACCCCTATTGTTTATTTTCTAAATACATTAATCAATATGATATC 3910  
Db 1864 GGGGAAATGTGCGGGAACCCCTATTGTTTATTTTCTAAATACATTAATCAATATGATATC 1923



3911 CGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAGAGTATGA 3970  
Db |||||  
1924 CGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAGAGTATGA 1983  
Qy |||||  
3971 GTATTCAACATTTCCGTTGTCGCTTATTCCTCTTTTGGGCAATTTTGGCTTCTCTCTTT 4030  
Db |||||  
1984 GTATTCAACATTTCCGTTGTCGCTTATTCCTCTTTTGGGCAATTTTGGCTTCTCTCTTT 2043  
Qy |||||  
4031 TTGCTCAACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGTGCAAG 4090  
Db |||||  
2044 TTGCTCAACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGTGCAAG 2103  
Qy |||||  
4091 TGGGTTACATCGAATCTGATCTCAACAGCGGTAGATCTTGGAGAGTTTGGCCCCGAAG 4150  
Db |||||  
2104 TGGGTTACATCGAATCTGATCTCAACAGCGGTAGATCTTGGAGAGTTTGGCCCCGAAG 2163  
Qy |||||  
4151 AACGTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATGTCGCGGTATTAATCCCGTA 4210  
Db |||||  
2164 AACGTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATGTCGCGGTATTAATCCCGTA 2223  
Qy |||||  
4211 TTGACGCGCGGCAAGAGCAACTCGGTCGCGCATACACTATTTCTCAGAATGACTTGTG 4270  
Db |||||  
2224 TTGACGCGCGGCAAGAGCAACTCGGTCGCGCATACACTATTTCTCAGAATGACTTGTG 2283  
Qy |||||  
4271 AGTACTCACAGTCAAGAAAGCATCTTACGAGTGCATGACAGTAAAGATTAATGCA 4330  
Db |||||  
2284 AGTACTCACAGTCAAGAAAGCATCTTACGAGTGCATGACAGTAAAGATTAATGCA 2343  
Qy |||||  
4331 GTCTGTCATTAACCATGAGTGAATTAACACTGCGGCAACTTACTTCTGACAAAGATCGAG 4390  
Db |||||  
2344 GTCTGTCATTAACCATGAGTGAATTAACACTGCGGCAACTTACTTCTGACAAAGATCGAG 2403  
Qy |||||  
4391 GACCGAAGGAGTAAACCGCTTTTGTGCAACATGCGGGATCATGTAACCTCGCTTGATC 4450  
Db |||||  
2404 GACCGAAGGAGTAAACCGCTTTTGTGCAACATGCGGGATCATGTAACCTCGCTTGATC 2463  
Qy |||||  
4451 GTTGGGAACCGGAGCTGAATGAGCCATACAAAGAGCGAGCGTGACACCAAGATGCTG 4510  
Db |||||  
2464 GTTGGGAACCGGAGCTGAATGAGCCATACAAAGAGCGAGCGTGACACCAAGATGCTG 2523  
Qy |||||  
4511 TAGCAATGCGCAACAACTGCGCAAACTATTAATCTGCGCAACTTACTTACTAGCTTCCC 4570  
Db |||||  
2524 TAGCAATGCGCAACAACTGCGCAAACTATTAATCTGCGCAACTTACTTACTAGCTTCCC 2583  
Qy |||||  
4571 GGCAACAATTAATAGACTGAGTGAAGCGGATAAAGTTGAGGACCACTTCTGCGCTCGG 4630  
Db |||||  
2584 GGCAACAATTAATAGACTGAGTGAAGCGGATAAAGTTGAGGACCACTTCTGCGCTCGG 2643  
Qy |||||  
4631 CCCTTCGGCTGGCTGTTTATTCCTGATTAATCTGGAGCGGCTGAGCGTGGTCTCGG 4690  
Db |||||  
2644 CCCTTCGGCTGGCTGTTTATTCCTGATTAATCTGGAGCGGCTGAGCGTGGTCTCGG 2703  
Qy |||||  
4691 GTATCATTTGAGCACTGGGCGCAGATGGTAGGCTCCCGTATGCTAGTTATCTACACGA 4750  
Db |||||  
2704 GTATCATTTGAGCACTGGGCGCAGATGGTAGGCTCCCGTATGCTAGTTATCTACACGA 2763  
Qy |||||  
4751 CGGGAGTCAGGCAACTATGATGAAGCAATAGACAGATCGCTGAGATAGTGCTCTCAC 4810  
Db |||||  
2764 CGGGAGTCAGGCAACTATGATGAAGCAATAGACAGATCGCTGAGATAGTGCTCTCAC 2823  
Qy |||||  
4811 TGATTAGCACTGGTACTGTCAGCAAGTTTACTCATATATCTTACTTATGATTAATTA 4870  
Db |||||  
2824 TGATTAGCACTGGTACTGTCAGCAAGTTTACTCATATATCTTACTTATGATTAATTA 2883  
Qy |||||  
4871 AACTTCATTTTAAATTAAGGATCTAGGTGAAGATCTTTTGTGATAATCTCATGACCA 4930  
Db |||||  
2884 AACTTCATTTTAAATTAAGGATCTAGGTGAAGATCTTTTGTGATAATCTCATGACCA 2943  
Qy |||||  
4931 AAATCCCTTAACTGAGTTTGGTTCCACTGAGCGGTGAGACCCCGGTGAGAAAGATCAAG 4990  
Db |||||  
2944 AAATCCCTTAACTGAGTTTGGTTCCACTGAGCGGTGAGACCCCGGTGAGAAAGATCAAG 3003  
Qy |||||  
4991 GATCTTCTGAGATCCCTTTTCTTCTGCGGCTAATCTCTGCTGTCGCAACAAAAAACCCAC 5050

3004 GATCTTCTGAGATCCCTTTTCTTCTGCGGTAATCTGCTGCTTGCACAAACAAAAAACCCAC 3063  
Qy |||||  
5051 CGCTTACAGGCGGTGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTTCGAGGTAA 5110  
Db |||||  
3064 CGCTTACAGGCGGTGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTTCGAGGTAA 3123  
Qy |||||  
5111 CTGGCTTACAGGCGGAGATACCAAAATCTGCTTCTAGTGTAGCGGTAGTTAGGCC 5170  
Db |||||  
3124 CTGGCTTACAGGCGGAGATACCAAAATCTGCTTCTAGTGTAGCGGTAGTTAGGCC 3183  
Qy |||||  
5171 ACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCTGCTCTAATCTCTTATCCAG 5230  
Db |||||  
3184 ACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCTGCTCTAATCTCTTATCCAG 3243  
Qy |||||  
5231 TGGCTGCTGCGAGTGGCGATAAGTCTGTTTACCGGGTTGGACTCAAGACGATAGTTAC 5290  
Db |||||  
3244 TGGCTGCTGCGAGTGGCGATAAGTCTGTTTACCGGGTTGGACTCAAGACGATAGTTAC 3303  
Qy |||||  
5291 CGGTAAGGCGCAGCGGTTCGGGCTGAACGGGGGTTCTGTGACACACAGCCGAGCTTGGAGC 5350  
Db |||||  
3304 CGGTAAGGCGCAGCGGTTCGGGCTGAACGGGGGTTCTGTGACACACAGCCGAGCTTGGAGC 3363  
Qy |||||  
5351 GAAAGACTTACACCGAACTGAGATACCTACAGCGTGAAGTATGAGAAAGCCGAGCTTTC 5410  
Db |||||  
3364 GAAAGACTTACACCGAACTGAGATACCTACAGCGTGAAGTATGAGAAAGCCGAGCTTTC 3423  
Qy |||||  
5411 CCAGAGGAGAGAAAGCGCGACAGGATATCCGGTAAAGCGGCGAGGTCGGAACAGGAGAGCGCA 5470  
Db |||||  
3424 CCAGAGGAGAGAAAGCGCGACAGGATATCCGGTAAAGCGGCGAGGTCGGAACAGGAGAGCGCA 3483  
Qy |||||  
5471 CGAGGAGAGTTCACAGGGGAAAACCGCTGCTGTTATCTTTATAGTCTCTGCTGGGTTTCCGACCC 5530  
Db |||||  
3484 CGAGGAGAGTTCACAGGGGAAAACCGCTGCTGTTATCTTTATAGTCTCTGCTGGGTTTCCGACCC 3543  
Qy |||||  
5531 TCTGACTTGAAGCTGCTGATTTTGTGATGCTGCTCAGGGGCGAGGCTATGGAAGAAACG 5590  
Db |||||  
3544 TCTGACTTGAAGCTGCTGATTTTGTGATGCTGCTCAGGGGCGAGGCTATGGAAGAAACG 3603  
Qy |||||  
5591 CCAGCAACCGCGCTTTTACGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5650  
Db |||||  
3604 CCAGCAACCGCGCTTTTACGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3663  
Qy |||||  
5651 TTCTGCTGTTATCCCTGATTTCTGTGATTAACCGTATTAACCGCTTCTGCTGCTGCTGCTGCTGCT 5710  
Db |||||  
3664 TTCTGCTGTTATCCCTGATTTCTGTGATTAACCGTATTAACCGCTTCTGCTGCTGCTGCTGCTGCT 3723  
Qy |||||  
5711 CCCTTCGCGCAGCGCAACCGAGCGCAGCGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 5770  
Db |||||  
3724 CCCTTCGCGCAGCGCAACCGAGCGCAGCGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 3783  
Qy |||||  
5771 GCCCAATACGCAACCGCTCTTCCCGCGGTTGGCGGATTCATTAATGCAAG 5822  
Db |||||  
3784 GCCCAATACGCAACCGCTCTTCCCGCGGTTGGCGGATTCATTAATGCAAG 3835

## RESULT 11

US-09-991-209-27  
; Sequence 27, Application US/09991209  
; Publication No. US20030024009A1  
; GENERAL INFORMATION:  
; APPLICANT: Dunn-Coleman, Nigel  
; APPLICANT: Langdon, Timothy  
; APPLICANT: Morse, Phillip  
; TITLE OF INVENTION: Manipulation of the Phenolic Acid  
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted  
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes  
; FILE REFERENCE: GC648-2  
; CURRENT APPLICATION NUMBER: US/09/991,209  
; PRIORITY FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: US 60/249,608  
; PRIORITY FILING DATE: 2000-11-17  
; NUMBER OF SEQ ID NOS: 97

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 5327
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pTP4a2 vector
US-09-991-209-27

Query Match          47.5%; Score 2768; DB 10; Length 5327;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 2895; Conservative 0; Mismatches 25; Indels 72; Gaps 4;

Qy 2835 AAGCAGATCGTTCAAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCTGTCGGGTC 2894
Db      |||
Qy 720  AAGCAGATCGTTCAAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCTGTCGGGTC 779
Db      |||
Qy 2895 TTGGATGATATATATATATTTCTGTGAATAGTTAGTAAAGCATGTAATTAACATGT 2954
Db      |||
Qy 780  TTGGATGATATATATATATTTCTGTGAATAGTTAGTAAAGCATGTAATTAACATGT 839
Db      |||
Qy 2955 AATCATGACGTTATTTATGATAGTGGTTTATGATTTAGAGTCCCGCAATTTATACATTT 3014
Db      |||
Qy 840  AATCATGACGTTATTTATGATAGTGGTTTATGATTTAGAGTCCCGCAATTTATACATTT 899
Db      |||
Qy 3015 AATACGCGATAGAAAAACAAATATAGCGCGCAAACTAGAGTAATAATTATCGCGCGGTGT 3074
Db      |||
Qy 900  AATACGCGATAGAAAAACAAATATAGCGCGCAAACTAGAGTAATAATTATCGCGCGGTGT 959
Db      |||
Qy 3075 CATCTATGTTACTAGATCGACCTCGAGGATGGGATCGCGCGCGCATGCGACGTGCGGC 3134
Db      |||
Qy 960  CATCTATGTTACTAGATCGATTAAGCTTCTAGAGCGCGGTGGAGC-----T 1006
Db      |||
Qy 3135 CCAATTCGCTATAGTAGTGTGTAATAC---AATTCACTGGCGGTGTTTACACGTC 3191
Db      |||
Qy 1007 CCAATTCGCTATAGTAGTGTGTAATACGGCGCTCACTGGCGGTGTTTACACGTC 1066
Db      |||
Qy 3192 GTGACTGGGAAAAACCCCTGGGCTTACCCAACTTAATCGCTTTGACGACATCCCGCTTTCG 3251
Db      |||
Qy 1067 GTGACTGGGAAAAACCCCTGGGCTTACCCAACTTAATCGCTTTGACGACATCCCGCTTTCG 1126
Db      |||
Qy 3252 CCAGCTGGGTAATAGGAGAGAGCCCGACCGATCGCCCTTCCCAACAGTTGCGGAGCC 3311
Db      |||
Qy 1127 CCAGCTGGGTAATAGGAGAGAGCCCGACCGATCGCCCTTCCCAACAGTTGCGGAGCC 1186
Db      |||
Qy 3312 TGAATGCGGAAT--GGAAGCGCTGTAGCGCGCATTAAGCGCGGCTGTGGTGGTTA 3370
Db      |||
Qy 1187 TGAATGCGGAATGGAAGCGCGCTGTAGCGCGCATTAAGCGCGGCTGTGGTGGTTA 1246
Db      |||
Qy 3371 CGCGCAGCGTGACCGCTACACTTCCAGCGCCCTAGCGCGCGCTCTTTCGCTTTCTTCC 3430
Db      |||
Qy 1247 CGCGCAGCGTGACCGCTACACTTCCAGCGCCCTAGCGCGCGCTCTTTCGCTTTCTTCC 1306
Db      |||
Qy 3431 CTTCTCTTTCGCGCAAGTGTGCGCGCTTTCGCGCTCAAGCTCTAAATCGGGGGCTCCCTT 3490
Db      |||
Qy 1307 CTTCTCTTTCGCGCAAGTGTGCGCGCTTTCGCGCTCAAGCTCTAAATCGGGGGCTCCCTT 1366
Db      |||
Qy 3491 TAGGGTTCCGATTTAGAGCTTTACGGCACCTTCGACCGCAAAACTTGAATTTGGGTGATG 3550
Db      |||
Qy 1367 TAGGGTTCCGATTTAGTGTCTTACGGCACCTTCGACCGCAAAACTTGAATTTAGGGTGA 1426
Db      |||
Qy 3551 GTTCACGTAGTGGGCCATCGCCCTGTAGAGAGGTTTTTCGCCCTTTGACGCTTGGAGTCCA 3610
Db      |||
Qy 1427 GTTCACGTAGTGGGCCATCGCCCTGTAGAGAGGTTTTTCGCCCTTTGACGCTTGGAGTCCA 1486
Db      |||
Qy 3611 CGTTCTTTAATAGTGAATCTGTGTTCCAACTGGAACAACTCAACCTCTATCTCGGTCT 3670
Db      |||
Qy 1487 CGTTCTTTAATAGTGAATCTGTGTTCCAACTGGAACAACTCAACCTCTATCTCGGTCT 1546
Db      |||
Qy 3671 ATTCTTTGATTTAAGGATTTTCGGCTATTCGGCTATTTGGTTAAAAAATGAGCTGA 3730
Db      |||
Qy 1547 ATTCTTTGATTTAAGGATTTTCGGCTATTCGGCTATTTGGTTAAAAAATGAGCTGA 1606
Db      |||

Qy 3731 TTTAAACAAATATTTAAACGCGAATTTTAAACAAATATTTAAACGTTTAAATTTCCCTGATG 3790
Db      |||
Qy 1607 TTTAAACAAATATTTAAACGCGAATTTTAAACAAATATTTAAACGTTTAAATTTCCCTGATG 1657
Db      |||
Qy 3791 CGGTATTTTCTCTTACGCACTCTGCGGTATTTACACACCGCATACAGTGCACCTTTTC 3850
Db      |||
Qy 1658 -----AGTGCGCACTTTTC 1671
Db      |||
Qy 3851 GGGGAAATGTGCGGGAACCCCTAATTTGTTTATTTTCTAAATACATTTCAAATATGTATC 3910
Db      |||
Qy 1672 GGGGAAATGTGCGGGAACCCCTAATTTGTTTATTTTCTAAATACATTTCAAATATGTATC 1731
Db      |||
Qy 3911 CGCTCATGAGACAATAAACCTTGATAAATGCTTCAATATATTTGAAAAAGGAAGATGTA 3970
Db      |||
Qy 1732 CGCTCATGAGACAATAAACCTTGATAAATGCTTCAATATATTTGAAAAAGGAAGATGTA 1791
Db      |||
Qy 3971 GTATTTCAACATTTCCGTTGCGCCCTTATTCCTTTTTCGCGCATTTTTCCTTCTGTTT 4030
Db      |||
Qy 1792 GTATTTCAACATTTCCGTTGCGCCCTTATTCCTTTTTCGCGCATTTTTCCTTCTGTTT 1851
Db      |||
Qy 4031 TTGCTCAACCCAGAAACGCTGCTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAG 4090
Db      |||
Qy 1852 TTGCTCAACCCAGAAACGCTGCTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAG 1911
Db      |||
Qy 4091 TGGGTTATCATGCAACTGATCTCAACAGCGGTAAAGATCTTTGAGAGTTTTCGCCCGGAG 4150
Db      |||
Qy 1912 TGGGTTATCATGCAACTGATCTCAACAGCGGTAAAGATCTTTGAGAGTTTTCGCCCGGAG 1971
Db      |||
Qy 4151 AACGTTTTCAAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGGTATTTATCCCGTA 4210
Db      |||
Qy 1972 AACGTTTTCAAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGGTATTTATCCCGTA 2031
Db      |||
Qy 4211 TTGACGCGGCGCAAGAGCAACTCGGTCGCGCATACACTATTTCTCAGAAATGACTTGGTTG 4270
Db      |||
Qy 2032 TTGACGCGGCGCAAGAGCAACTCGGTCGCGCATACACTATTTCTCAGAAATGACTTGGTTG 2091
Db      |||
Qy 4271 AGTACTCAGAGTCAAGAAAGCATCTTACGATGGCATGACAGTAAGAGAAATATGCA 4330
Db      |||
Qy 2092 AGTACTCAGAGTCAAGAAAGCATCTTACGATGGCATGACAGTAAGAGAAATATGCA 2151
Db      |||
Qy 4331 GTGCTGCATTAACCATGAGTGAATAACACTGCGGCAACTTACTTCTGACACAGATCGGAG 4390
Db      |||
Qy 2152 GTGCTGCATTAACCATGAGTGAATAACACTGCGGCAACTTACTTCTGACACAGATCGGAG 2211
Db      |||
Qy 4391 GACCGAAGGAGCTAAACCGCTTTTTCACAAACATGCGGCGATCATGTAACCTGCGCTGATC 4450
Db      |||
Qy 2212 GACCGAAGGAGCTAAACCGCTTTTTCACAAACATGCGGCGATCATGTAACCTGCGCTGATC 2271
Db      |||
Qy 4451 GTTGGGAAACCGGAGCTGAATGAAGCCATACCAAACGAGCGGTGACACAGATGCGCTG 4510
Db      |||
Qy 2272 GTTGGGAAACCGGAGCTGAATGAAGCCATACCAAACGAGCGGTGACACAGATGCGCTG 2331
Db      |||
Qy 4511 TAGCAATGGCAACCAACCTGCGCAAACTATTAACCTGCGCAACTTACTTCTAGCTTCCC 4570
Db      |||
Qy 2332 TAGCAATGGCAACCAACCTGCGCAAACTATTAACCTGCGCAACTTACTTCTAGCTTCCC 2391
Db      |||
Qy 4571 GGCACAAATTAATAGACTGGAATGAGGCGGATTAAGTTTGCAGGACCACTTCTTGCCTCGG 4630
Db      |||
Qy 2392 GGCACAAATTAATAGACTGGAATGAGGCGGATTAAGTTTGCAGGACCACTTCTTGCCTCGG 2451
Db      |||
Qy 4631 CCTTTCGCGCTGCGTGTGTTTATTCCTGATTAATCTGGAGCGGTGAGCGGTGCGGCTGCGG 4690
Db      |||
Qy 2452 CCTTTCGCGCTGCGTGTGTTTATTCCTGATTAATCTGGAGCGGTGAGCGGTGCGGCTGCGG 2511
Db      |||
Qy 4691 GTATCATTTGAGCACTGCGGCGCATGTAAGCCCTCCGCTATCTGTTATCTACACGA 4750
Db      |||
Qy 2512 GTATCATTTGAGCACTGCGGCGCATGTAAGCCCTCCGCTATCTGTTATCTACACGA 2571
Db      |||
Qy 4751 CGGGAGTTCAGGCAACTTATGATGAACGAATAGACAGATCGCTGAGATAGGTGCTCAC 4810
Db      |||
Qy 2572 CGGGAGTTCAGGCAACTTATGATGAACGAATAGACAGATCGCTGAGATAGGTGCTCAC 2631
Db      |||
Qy 4811 TGATTAAGCATTTGGTAACTGTCTCAGACCAAGTTTACTCATATATATCTTTAGATTGATTAA 4870
Db      |||
```



```
Db 2632 TGAATAGCAATGTAAGTCTGACAGCAGTTTACTCATATATACCTTAGATGATTAA 2691
Qy 4871 AACTTCATTTTAAATTAAGAGTCTAGGTGAAGTCTTTTGTGTAATCTCATGACCA 4930
Db 2692 AACTTCATTTTAAATTAAGAGTCTAGGTGAAGTCTTTTGTGTAATCTCATGACCA 2751
Qy 4931 AAATCCCTTAAGCTGAGTTTCTGTCACCTGAGGCTCAGACCCGTAGAAAAGATCAAG 4990
Db 2752 AAATCCCTTAAGCTGAGTTTCTGTCACCTGAGGCTCAGACCCGTAGAAAAGATCAAG 2811
Qy 4991 GATCTCTCTGAGATCCCTTTTCTGCGCGTAATCTGCTGCTGCAACCAAAAAACCA 5050
Db 2812 GATCTCTCTGAGATCCCTTTTCTGCGCGTAATCTGCTGCTGCAACCAAAAAACCA 2871
Qy 5051 CGCTACAGCGGTGTTGTTGCGCGATCAAGAGCTACCAACTCTTTTCCGAAGTAA 5110
Db 2872 CGCTACAGCGGTGTTGTTGCGCGATCAAGAGCTACCAACTCTTTTCCGAAGTAA 2931
Qy 5111 CTGGCTTCAGAGCGCAGATACCAAAATCTGCTCTCTAGTGTAGCCGTAGTAGGCC 5170
Db 2932 CTGGCTTCAGAGCGCAGATACCAAAATCTGCTCTCTAGTGTAGCCGTAGTAGGCC 2991
Qy 5171 ACCACTCAAGAACTCTGTAGCAGCGCTACATACCTCTGCTCTGCTAACTCTGTTACCAG 5230
Db 2992 ACCACTCAAGAACTCTGTAGCAGCGCTACATACCTCTGCTCTGCTAACTCTGTTACCAG 3051
Qy 5231 TGGCTGCTGCGAGTGGCGATAAGTCTGCTCTACCGGGTGGAGCTCAAGACGATAGTTAC 5290
Db 3052 TGGCTGCTGCGAGTGGCGATAAGTCTGCTCTACCGGGTGGAGCTCAAGACGATAGTTAC 3111
Qy 5291 CGGATAAGCGCGAGCGGTGCGGCTGAACCGGGGGTTCGTGCAACAGCCAGCTTGAGC 5350
Db 3112 CGGATAAGCGCGAGCGGTGCGGCTGAACCGGGGGTTCGTGCAACAGCCAGCTTGAGC 3171
Qy 5351 GACGACCTTACCGAACTCAGATACCTACAGCGTGGCTATCAGAAAGCGCCACGCTTC 5410
Db 3172 GACGACCTTACCGAACTCAGATACCTACAGCGTGGCTATCAGAAAGCGCCACGCTTC 3231
Qy 5411 CCGAAGGAGAAAGCGGAGCAGGTATCCGGTAAGCGCGCAGGGTCGAAACAGAGAGCGCA 5470
Db 3232 CCGAAGGAGAAAGCGGAGCAGGTATCCGGTAAGCGCGCAGGGTCGAAACAGAGAGCGCA 3291
Qy 5471 CGAGGAGCTTCAGGGGAGAAACGCTCTGTATCTTTATAGTCTCTGCGGGTTTCGCCACC 5530
Db 3292 CGAGGAGCTTCAGGGGAGAAACGCTCTGTATCTTTATAGTCTCTGCGGGTTTCGCCACC 3351
Qy 5531 TCTGACTTGAGCGTTCGATTTTCTGATGCTCTGTCAGGGGGCGGAGCTATGCAAAACG 5590
Db 3352 TCTGACTTGAGCGTTCGATTTTCTGATGCTCTGTCAGGGGGCGGAGCTATGCAAAACG 3411
Qy 5591 CCAGCAACCGCGCTTTTACGGTCTCTGCGCTTTTCTGCGCTTTTGTCTCAGATGTTCT 5650
Db 3412 CCAGCAACCGCGCTTTTACGGTCTCTGCGCTTTTGTCTCAGATGTTCT 3471
Qy 5651 TTCTCTGCTTATCCCTGATTTCTGTGATAACCGTATTAACCGCTTTGAGTGAGCTGATA 5710
Db 3472 TTCTCTGCTTATCCCTGATTTCTGTGATAACCGTATTAACCGCTTTGAGTGAGCTGATA 3531
Qy 5711 CGGCTCGCGCAGCGCAACGACCGAGCGAGCTCAGTGACGAGGAGCGGAGAGC 5770
Db 3532 CGGCTCGCGCAGCGCAACGACCGAGCGAGCTCAGTGACGAGGAGCGGAGAGC 5822
Qy 5771 GCCCAATACGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGACG 5822
Db 3592 GCCCAATACGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGACG 3643
```

RESULT 12

US-09-991-209-19

; Sequence 19, Application US/09991209

; Publication No. US20030024009A1

; GENERAL INFORMATION:

```
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Langdon, Timothy
; APPLICANT: Morse, Phillip
; TITLE OF INVENTION: Manipulation of the Phenolic Acid
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes
; FILE REFERENCE: GC648-2
; CURRENT APPLICATION NUMBER: US/09/991,209
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/249,608
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 5337
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pTU4 vector
; US-09-991-209-19
```

```
Query Match 47.5%; Score 2768; DB 10; Length 5337;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 2895; Conservative 0; Mismatches 25; Indels 72; Gaps 4;
```

```
Qy 2835 AAGCAGATCGTTCAAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCTGTTCGCGGTC 2894
Db 954 AAGCAGATCGTTCAAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCTGTTCGCGGTC 1013
Qy 2895 TTGCGGATGATTATCATATAATTTCTGTTGAATTAAGTTAAGCATGTAATAATTAACATGT 2954
Db 1014 TTGCGGATGATTATCATATAATTTCTGTTGAATTAAGTTAAGCATGTAATAATTAACATGT 1073
Qy 2955 AATGATGATGATTTATTTATGAGATGGTGTATTTATGATTAAGTATCCGCAATTAACATTT 3014
Db 1074 AATGATGATGATTTATTTATGAGATGGTGTATTTATGATTAAGTATCCGCAATTAACATTT 1133
Qy 3015 AATACCGGATAGAAAACAAATATAGCGCGCAAACTTAGGATAAATTTATCGCGCGGTGT 3074
Db 1134 AATACCGGATAGAAAACAAATATAGCGCGCAAACTTAGGATAAATTTATCGCGCGGTGT 1193
Qy 3075 CATCTATGTTACTAGATCGACCTGACAGGATGGGATCCGCGCGCATGCGACGTGCGGC 3134
Db 1194 CATCTATGTTACTAGATCGATTAAGCTTCTAGAGCGCGGTGGAGC-----T 1240
Qy 3135 CCAATTGCGCCCTATAGTGAATCGTATTAC---AATTCACTGGCGCGTGTTTTACAACGTC 3191
Db 1241 CCAATTGCGCCCTATAGTGAATCGTATTACCGCGCGCTCACTGGCGCGTGTTTTACAACGTC 1300
Qy 3192 GTGACTGGGAAACCCCTGGCGTTACCCAACTTAATCGCTTTCAGCAGCAGCATCCCGCTTTCG 3251
Db 1301 GTGACTGGGAAACCCCTGGCGTTACCCAACTTAATCGCTTTCAGCAGCAGCATCCCGCTTTCG 1360
Qy 3252 CCAGCTGGGATTAAGCGAAGGCGCGCACCGATCGCCCTTCCCAACAGTTCGGAGCC 3311
Db 1361 CCAGCTGGGATTAAGCGAAGGCGCGCACCGATCGCCCTTCCCAACAGTTCGGAGCC 1420
Qy 3312 TGAATGGCGCAAT--GGACGCGCTGTAGCGGCGCATTAAGCGCGCGGTGTGTGTGTTA 3370
Db 1421 TGAATGGCGCAATGGAGACGCGCTGTAGCGGCGCATTAAGCGCGCGGTGTGTGTGTTA 1480
Qy 3371 CGCGCAGCGTGAACCGCTACACTTTGCGAGCGCTTACGCGCGCTTTCCTTTCTTCTTCC 3430
Db 1481 CGCGCAGCGTGAACCGCTACACTTTGCGAGCGCTTACGCGCGCTTTCCTTTCTTCTTCC 1540
Qy 3431 CTTCCTTTCTCGCAGCTTTCGCGGCTTTCGCGCTCAAGCTCTAAATCGGGGCTCCCTT 3490
Db 1541 CTTCCTTTCTCGCAGCTTTCGCGGCTTTCGCGCTCAAGCTCTAAATCGGGGCTCCCTT 1600
Qy 3491 TAGGGTTCCGATTTAGAGCTTTTACGCGCATCTCCACCGCAAAACCTTGAATTGGGTGATG 3550
Db 1601 TAGGGTTCCGATTTAGTGTCTTTACGCGCATCTCCACCGCAAAACCTTGAATTGGGTGATG 1660
```

Qy	3551	GTTCACGTAGTGGCCATCCGCCCTGTATAGACGGTCTTTTCGCCCTTTGACGCTTGGAGTCCA	3610		Db	2686	CCCTTCGGCTGGCTGGTCTTTATTTGCTGATAAATCTGGAGCCGGTGAGCGTGGCTCTCCGG	2745	
Db	1661	GTTCACGTAGTGGCCATCCGCCCTGTATAGACGGTCTTTTCGCCCTTTGACGCTTGGAGTCCA	1720		Qy	4691	GTATCATTTGACGACACTGGGGCCAGATGGTAAGCCCTCCCGCTATCGTAGTATCTACACGA	4750	
Qy	3611	CGTCTCTTAATAGTGGACTCTTGGTCTCABAACTGGAACAACACTCAACCCCTATCTGGCTCT	3670		Db	2746	GTATCATTTGACGACACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTATCTACACGA	2805	
Db	1721	CGTCTCTTAATAGTGGACTCTTGGTCTCABAACTGGAACAACACTCAACCCCTATCTGGCTCT	1780		Qy	4751	CGGGAGCTCAGGCAACTATGSAATAGAACGAAATAGACAGATCGCTGAGATAGGTGCTCAC	4810	
Qy	3671	ATTCTTTTGATTTATAAGGGAATTTTGGCGATTTTCGGCTATTTGGTTAAATAATGAGCTGA	3730		Db	2806	CGGGAGCTCAGGCAACTATGSAATAGAACGAAATAGACAGATCGCTGAGATAGGTGCTCAC	2865	
Db	1781	ATTCTTTTGATTTATAAGGGAATTTTGGCGATTTTCGGCTATTTGGTTAAATAATGAGCTGA	1840		Qy	4811	TGATTAAGCATTTGTAATCTGTACAGACCAAGTTTACTCATATATATCTTTAGATTTAA	4870	
Qy	3731	TTTAAACAAATATTTAAACGCAATTTTAAACAAATATTTAAACGTTTACAAATTTGCGCTGATG	3790		Db	2866	TGATTAAGCATTTGTAATCTGTACAGACCAAGTTTACTCATATATATCTTTAGATTTAA	2925	
Db	1841	TTTAAACAAATATTTAAACGCAATTTTAAACAAATATTTAAACGTTTACAAATTTT	1891		Qy	4871	AACCTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCTTTTGTGATAATCTCATGACCA	4930	
Qy	3791	CGGTATTTTCTCCTTACGCATCTGTGCGGTATTTTACACGCCATACAGGTGGCACTTTTC	3850		Db	2926	AACCTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCTTTTGTGATAATCTCATGACCA	2985	
Db	1892	-----AGGTGGCACTTTTC-----	1905		Qy	4931	AAATCCCTTAAACGTGAGTTTTCCTTCACTGAGCGTCAAGCCCGTAGAAGATCAAG	4990	
Qy	3851	GGGGAATGTGCGGGAAACCCCTATTTGTTTATTTTCTAAATACATTTCAAATATGTATC	3910		Db	2986	AAATCCCTTAAACGTGAGTTTTCCTTCACTGAGCGTCAAGCCCGTAGAAGATCAAG	3045	
Db	1906	GGGGAATGTGCGGGAAACCCCTATTTGTTTATTTTCTAAATACATTTCAAATATGTATC	1965		Qy	4991	GATCTTCTTGAGATCTTTTCTGCGCGTAATCTGCTGCTTGCATAAACAACCAACAC	5050	
Qy	3911	CGCTCATGAGACAATAACCCCTGATAAATGCTTCAATATATTTGAAAGGAAGATATGA	3970		Db	3046	GATCTTCTTGAGATCTTTTCTGCGCGTAATCTGCTGCTTGCATAAACAACCAACAC	3105	
Db	1966	CGCTCATGAGACAATAACCCCTGATAAATGCTTCAATATATTTGAAAGGAAGATATGA	2025		Qy	5051	CGCTTACAGCGGTGTTTGTGTCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAA	5110	
Qy	3971	GTATTCACATTTTCGGTGTGCGCTTATTCCTTTTGTGGCGATTTTGGCTTCCCTGTTT	4030		Db	3106	CGCTTACAGCGGTGTTTGTGTCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAA	3165	
Db	2026	GTATTCACATTTTCGGTGTGCGCTTATTCCTTTTGTGGCGATTTTGGCTTCCCTGTTT	2085		Qy	5111	CTGCTTTCAGCAGCGCGAGATACCAATATCTGCTCTTCTAGTGTAGCGGTAGTAGGCC	5170	
Qy	4031	TTGCTCACCCAGAAACGCTGTGTAAGTAAAGATGCTGAAAGATCAGTTGGGTGCAAG	4090		Db	3166	CTGCTTTCAGCAGCGCGAGATACCAATATCTGCTCTTCTAGTGTAGCGGTAGTAGGCC	3225	
Db	2086	TTGCTCACCCAGAAACGCTGTGTAAGTAAAGATGCTGAAAGATCAGTTGGGTGCAAG	2145		Qy	5171	ACCACTTCAAGAACTCTGTAGTACACCGCTTACATACCTCGCTCTGCTATCTGTTTACCAG	5230	
Qy	4091	TGGGTTCATCGAATCTGATCTCAACAGCGGTAGATCTTTGAGAGTTTTCGCCCGGAAG	4150		Db	3226	ACCACTTCAAGAACTCTGTGTAGTACACCGCTTACATACCTCGCTCTGCTATCTGTTTACCAG	3285	
Db	2146	TGGGTTCATCGAATCTGATCTCAACAGCGGTAGATCTTTGAGAGTTTTCGCCCGGAAG	2205		Qy	5231	TGGCTGTCTCCAGTGGCGGATAAAGTCTGTCTTCTTACCGGTTTGGACTCAAGAGCGATAGTTAC	5290	
Qy	4151	AACGTTTTCGAATGAGCACTTTTAAAGTTCTGCTATGTGGCGGTATTTATCCCGTA	4210		Db	3286	TGGCTGTCTCCAGTGGCGGATAAAGTCTGTCTTCTTACCGGTTTGGACTCAAGAGCGATAGTTAC	3345	
Db	2206	AACGTTTTCGAATGAGCACTTTTAAAGTTCTGCTATGTGGCGGTATTTATCCCGTA	2265		Qy	5291	CGGATAAGCGCGACGCGTCTGGGCTGAACCGGGGTTCTGTGACACAGCCGCTTGGAGC	5350	
Qy	4211	TTGACGCGGCGCAAGAGCAACTCGGTGCGCGCATACATTTCTCAGAAATGACTTGGTTG	4270		Db	3346	CGGATAAGCGCGACGCGTCTGGGCTGAACCGGGGTTCTGTGACACAGCCGCTTGGAGC	3405	
Db	2266	TTGACGCGGCGCAAGAGCAACTCGGTGCGCGCATACATTTCTCAGAAATGACTTGGTTG	2325		Qy	5351	GAAACGCTACACCGAACTGAGATACCTTACAGCGGTGAGCTATGAGAAAGCGCACGCTTC	5410	
Qy	4271	AGTACTCACAGTCAAGAAAGCATTTTACGGATGGCATGACAGTAAAGAAATATGCA	4330		Db	3406	GAAACGCTACACCGAACTGAGATACCTTACAGCGGTGAGCTATGAGAAAGCGCACGCTTC	3465	
Db	2326	AGTACTCACAGTCAAGAAAGCATTTTACGGATGGCATGACAGTAAAGAAATATGCA	2385		Qy	5411	CCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGCGGAGGTTCGGAACAGGAGAGCGCA	5470	
Qy	4331	GTGCTGCCATACCATGAGTATTAACACTGCGGCCAATTTACTTCTGACACGATCGGAG	4390		Db	3466	CCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGCGGAGGTTCGGAACAGGAGAGCGCA	3525	
Db	2386	GTGCTGCCATACCATGAGTATTAACACTGCGGCCAATTTACTTCTGACACGATCGGAG	2445		Qy	5471	CGAGGAGCTTTCAGGGGGAACCGCTGGTATCTTTTATAGTCTCTGCGGTTTCCGACC	5530	
Qy	4391	GACCGAAGGAGTAAACCGCTTTTTCACAAATGCGGGGATCATGTAACTCGCCTTGATC	4450		Db	3526	CGAGGAGCTTTCAGGGGGAACCGCTGGTATCTTTTATAGTCTCTGCGGTTTCCGACC	3585	
Db	2446	GACCGAAGGAGTAAACCGCTTTTTCACAAATGCGGGGATCATGTAACTCGCCTTGATC	2505		Qy	5531	TCTGACTTGAAGCTGATTTTGTGATGCTGCTGAGGGGGCGGAGCTTATGGAAGAAAGCG	5590	
Qy	4451	GTGCGGAAACCGGAGCTGAATGAAGCCATACAAACGACGAGGTGACACCGATGCGCTG	4510		Db	3586	TCTGACTTGAAGCTGATTTTGTGATGCTGCTGAGGGGGCGGAGCTTATGGAAGAAAGCG	3645	
Db	2506	GTGCGGAAACCGGAGCTGAATGAAGCCATACAAACGACGAGGTGACACCGATGCGCTG	2565		Qy	5591	CCAGCAACCGCGCTTTTAAAGCTTCTGCGCTTTTGTGCGCTTTTGTCTCATATGTTCT	5650	
Qy	4511	TAGCAATGCCAACACGTTGCGCAACTTTTAACTGCGGCACTTACTTACTTACTGCTTCCC	4570		Db	3646	CCAGCAACCGCGCTTTTAAAGCTTCTGCGCTTTTGTGCGCTTTTGTCTCATATGTTCT	3705	
Db	2566	TAGCAATGCCAACACGTTGCGCAACTTTTAACTGCGGCACTTACTTACTTACTGCTTCCC	2625		Qy	5651	TTCTGCTTATTCCTGATTTCTGTGGAATAACCGTATTTACCGCTTTTGTGAGTGTGATTA	5710	
Qy	4571	GCCAAATTTAATGACTGAGTGGAGCGGATTAAGTTTGCAGGACCACTTCTGCGCTCGG	4630		Db	3706	TTCTGCTTATTCCTGATTTCTGTGGAATAACCGTATTTACCGCTTTTGTGAGTGTGATTA	3765	
Db	2626	GCCAAATTTAATGACTGAGTGGAGCGGATTAAGTTTGCAGGACCACTTCTGCGCTCGG	2685		Qy	5711	CCGCTCGCGCGCAGCGCAACGCGAGCGAGCTCAGTGTAGCGGAGCGGAGGAGC	5770	
Qy	4631	CCCTTCGCGCTGGCTGTTTATTTGCTGATTAATCTGGAGCGGTGAGCGGTGCTCGCG	4690						

Db 3766 CCCTCGCCGACGCCGAAACGACGAGCGAGCGAGTCAGTGCAGGAGGAAGCGAAGAGC 3825  
Qy 5771 GCCCAATACGCAACCGCTCTCCCGCGCGTGTGGCCGATTCATTAATGCAG 5822  
Db 3826 GCCCAATACGCAACCGCTCTCCCGCGCGTGTGGCCGATTCATTAATGCAG 3877

RESULT 13  
US-09-991-209-23  
; Sequence 23, Application US/09991209  
; Publication No. US20030024009A1  
; GENERAL INFORMATION:  
; APPLICANT: Dunn-Coleman, Nigel  
; APPLICANT: Langdon, Timothy  
; APPLICANT: Morse, Phillip  
; TITLE OF INVENTION: Manipulation of the Phenolic Acid  
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted  
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes  
; FILE REFERENCE: GC648-2  
; CURRENT APPLICATION NUMBER: US/09/991,209  
; PRIOR FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: US 60/249,608  
; PRIOR FILING DATE: 2000-11-17  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 5337  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pTP8-5 vector  
US-09-991-209-23

Query Match 47.5%; Score 2768; DB 10; Length 5337;  
Best Local Similarity 96.8%; Pred. No. 0;  
Matches 2895; Conservative 0; Mismatches 25; Indels 72; Gaps 4;

Qy 2835 AAGCAGATCGTTCCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATTCCTGTTCGGTC 2894  
Db 954 AAGCAGATCGTTCCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATTCCTGTTCGGTC 1013

Qy 2895 TTGGGATGATTCATATAATTTCTGTGAATTAAGTGAAGCATGTAATAATTAACATGT 2954  
Db 1014 TTGGGATGATTCATATAATTTCTGTGAATTAAGTGAAGCATGTAATAATTAACATGT 1073

Qy 2955 AATGCAATGATTAATTAAGATGGTGTTCATGATTAAGTTCGCGCAATTAACATTT 3014  
Db 1074 AATGCAATGATTAATTAAGATGGTGTTCATGATTAAGTTCGCGCAATTAACATTT 1133

Qy 3015 AATACCGGATAGAAAACAAATAATAGCGCAAACTAGGATAAATTAATTCGCGCGGTGT 3074  
Db 1134 AATACCGGATAGAAAACAAATAATAGCGCAAACTAGGATAAATTAATTCGCGCGGTGT 1193

Qy 3075 CATCTATGTTACTAGATCGACCTGCGAGCATGGATTCGCGCGCGCATGCGAGCTCGGC 3134  
Db 1194 CATCTATGTTACTAGATCGACCTGCGAGCATGGATTCGCGCGCGCATGCGAGCTCGGC 1240

Qy 3135 CCAATTCGCGCTTACTAGTGTGCTATTAC --- AATTCACTGGCGGTGTTTACAACGTC 3191  
Db 1241 CCAATTCGCGCTTACTAGTGTGCTATTACCGCGCTCACTGGCGGTGTTTACAACGTC 1300

Qy 3192 GTGACTGGGAAAACCTGGCGTTACCCAACTTAATTCGCTTGCAGCACATPCCCTTTCG 3251  
Db 1301 GTGACTGGGAAAACCTGGCGTTACCCAACTTAATTCGCTTGCAGCACATPCCCTTTCG 1360

Qy 3252 CCAGCTGGGTATAGGAGAGCGCGACCGATGCCCTTCCACAGTTGCGGAGCC 3311  
Db 1361 CCAGCTGGGTATAGGAGAGCGCGACCGATGCCCTTCCACAGTTGCGGAGCC 1420

Qy 3312 TGAATGCGCAAT -GGACGCGCCCTGTAGCGCGCATTAAGCGCGCGGTGTGCTGTTA 3370  
Db 1421 TGAATGCGCAATGGAAGCGCGCCCTGTAGCGCGCATTAAGCGCGCGGTGTGCTGTTA 1480

Qy 3371 CGCGAGCGTACCGCTACACTTGCAGCGCCCTAGCGCCCGCTCTCTTTCGCTTCTTCC 3430  
Db 1481 CGCGAGCGTACCGCTACACTTGCAGCGCCCTAGCGCCCGCTCTCTTTCGCTTCTTCC 1540

Qy 3431 CTTTCCTTTCTCGCCACGTTTCGCCGCTTTCGCCGCTCAAGCTCTAATAATCGGGGCTCCCTT 3490  
Db 1541 CTTTCCTTTCTCGCCACGTTTCGCCGCTTTCGCCGCTCAAGCTCTAATAATCGGGGCTCCCTT 1600

Qy 3491 TAGGGTTCGATTTAGAGCTTTTACGACCTTCGACCGCAAAACAACTTGTATGGGTGATG 3550  
Db 1601 TAGGGTTCGATTTAGAGCTTTTACGACCTTCGACCGCAAAACAACTTGTATGGGTGATG 1660

Qy 3551 GTTTCAGTAGTGGCCCATCGCCCTGATAGACGTTTTTTCGCCCTTTCGACGTTGAGTCCA 3610  
Db 1661 GTTTCAGTAGTGGCCCATCGCCCTGATAGACGTTTTTTCGCCCTTTCGACGTTGAGTCCA 1720

Qy 3611 CGTTCTTTAATAGTACGACTCTGTCTCAAACTGCAAACTCAACCTATCTCGGTCT 3670  
Db 1721 CGTTCTTTAATAGTACGACTCTGTCTCAAACTGCAAACTCAACCTATCTCGGTCT 1780

Qy 3671 ATTCTTTTGAATTTAAGGGAATTTTGCAGATTTTGCCTATTTGGTTAAAAAATGAGCTGA 3730  
Db 1781 ATTCTTTTGAATTTAAGGGAATTTTGCAGATTTTGCCTATTTGGTTAAAAAATGAGCTGA 1840

Qy 3731 TTAAACAATATTTAAACGGAATTTTAAACAATATTAAGTTTACATTTTCGCTGATG 3790  
Db 1841 TTAAACAATATTTAAACGGAATTTTAAACAATATTAAGTTTACATTTTTCGCTGATG 1891

Qy 3791 CGGTATTTTCTCTTACGATCTGTGGGTATTTTACACCGCATACAGGTGGCACTTTTC 3850  
Db 1892 -----AGGTGGCACTTTTC 1905

Qy 3851 GGCGAATGTGCGGGAACCCCTATTTTGTATTTTCTAAATACATTTCAAAATATGATTC 3910  
Db 1906 GGCGAATGTGCGGGAACCCCTATTTTGTATTTTCTAAATACATTTCAAAATATGATTC 1965

Qy 3911 CGCTCATGAGACATAAACCCTGATAAATGCTTCAATAATTAAGAAAGAAAGATGATGA 3970  
Db 1966 CGCTCATGAGACATAAACCCTGATAAATGCTTCAATAATTAAGAAAGAAAGATGATGA 2025

Qy 3971 GTATTCAACATTTCCGCTGCGCTTATTTCCCTTTTTCGCGCATTTTTCCTTCTGTTT 4030  
Db 2026 GTATTCAACATTTCCGCTGCGCTTATTTCCCTTTTTCGCGCATTTTTCCTTCTGTTT 2085

Qy 4031 TTGCTCACCCAGAAAACGCTGGTGAAGATGCTGAAGATCAGTTGGGTGCGAGAG 4090  
Db 2086 TTGCTCACCCAGAAAACGCTGGTGAAGATGCTGAAGATCAGTTGGGTGCGAGAG 2145

Qy 4091 TGGGTTTACATCGAATGATCTCAACAGCGGTAAAGATCCTTGAAGATTTTTCGCCCGAAG 4150  
Db 2146 TGGGTTTACATCGAATGATCTCAACAGCGGTAAAGATCCTTGAAGATTTTTCGCCCGAAG 2205

Qy 4151 AAGTTTTTCCAATGATGAGCATTTTTAAAGTTCCTGCTATGTGCGCGGTATTTATCCGTA 4210  
Db 2206 AAGTTTTTCCAATGATGAGCATTTTTAAAGTTCCTGCTATGTGCGCGGTATTTATCCGTA 2265

Qy 4211 TTGAOCGCGGCAAGAGCAACTCGGTGCGCGCATACATTTCTCAGAATGACTTGGTGG 4270  
Db 2266 TTGAOCGCGGCAAGAGCAACTCGGTGCGCGCATACATTTCTCAGAATGACTTGGTGG 2325

Qy 4271 AGTACTCACCGGTTCAGAAAGCATCTTACGATGCGATGACAGTAAGAGAAATTAATGCA 4330  
Db 2326 AGTACTCACCGGTTCAGAAAGCATCTTACGATGCGATGACAGTAAGAGAAATTAATGCA 2385

Qy 4331 GTGCTGCCATAACCATGAGTGAATACTGCGGCCAACTTACTTCTGACAAACGATCGGAG 4390  
Db 2386 GTGCTGCCATAACCATGAGTGAATACTGCGGCCAACTTACTTCTGACAAACGATCGGAG 2445

Qy 4391 GACCGAAGAGGCTAAACCGCTTTTTCGCAAACTATGGGGGATCATGTAACCTGCGCTGATC 4450  
Db 2446 GACCGAAGAGGCTAAACCGCTTTTTCGCAAACTATGGGGGATCATGTAACCTGCGCTGATC 2505

Qy 4451 GTTGGGAAACCGGAGCTGAATGAAGCCATACCAACGAGCGGTGACACCAAGATGCTG 4510



Qy 3192 GTGACTGGGAAACCCCTGGCGTTACCAACTTAATCGCTTGACGACATACCCCTTTTCG 3251  
Db GTGACTGGGAAACCCCTGGCGTTACCAACTTAATCGCTTGACGACATACCCCTTTTCG 1360  
Qy 3252 CAGCTGGCGTAATAGCGAGAGAGCCCGACCGATCGCCTTCCCAACAGTTGCGGAGCC 3311  
Db CAGCTGGCGTAATAGCGAGAGAGCCCGACCGATCGCCTTCCCAACAGTTGCGGAGCC 1420  
Qy 3312 TGAATGGCGAAT -CGACGCGCCCTGTAGCGCGCATTAAGCGCGCGGCTGTGCGTGA 3370  
Db TGAATGGCGAATGGAGCGCGCCCTGTAGCGCGCATTAAGCGCGCGGCTGTGCGTGA 1480  
Qy 3371 CGCGCAGCGTGACCGCTACACTTCGCCAGCGCCCTAGCGCCCGCTCTCTTCCTTCTTC 3430  
Db CGCGCAGCGTGACCGCTACACTTCGCCAGCGCCCTAGCGCCCGCTCTCTTCCTTCTTC 1540  
Qy 3431 CTTCTCTTTCGCCACAGTTTCCCGGCTTCCCGCTCAAGCTCTAAATCGGCGGCTCCCTT 3490  
Db CTTCTCTTTCGCCACAGTTTCCCGGCTTCCCGCTCAAGCTCTAAATCGGCGGCTCCCTT 1600  
Qy 3491 TAGGGTTCCGATTTAGAGCTTTAGCGCACCTCGACCGCAAAACTTGAATTTGGTGTATG 3550  
Db TAGGGTTCCGATTTAGAGCTTTAGCGCACCTCGACCGCAAAACTTGAATTTGGTGTATG 1660  
Qy 3551 GTTCACTAGTGGGCCATCGCCCTGATAGAGGTTTTTTCGCCCTTTGACGCTTGAGTCCA 3610  
Db GTTCACTAGTGGGCCATCGCCCTGATAGAGGTTTTTTCGCCCTTTGACGCTTGAGTCCA 1720  
Qy 3611 CGTCTTTAATAGTGGACTCTGTTCCAAACTGGAACCAACTCAACCTTATCTCGGTCT 3670  
Db CGTCTTTAATAGTGGACTCTGTTCCAAACTGGAACCAACTCAACCTTATCTCGGTCT 1780  
Qy 3671 ATTCTTTTGAATTAAGGGAATTTTTCGGCTTTCGGCTTATTTGTTAAATAATGAGCTGA 3730  
Db ATTCTTTTGAATTAAGGGAATTTTTCGGCTTTCGGCTTATTTGTTAAATAATGAGCTGA 1840  
Qy 3731 TTTAAACAAATTAATTAACCGCAATTTTAAACAAATTAATTAACCGCTTACAAATTTTCG 3790  
Db TTTAAACAAATTAATTAACCGCAATTTTAAACAAATTAATTAACCGCTTACAAATTT 1891  
Qy 3791 CGGTATTTTCTCTACGATCTGCGGATTTTTCACACCGCATACAGGTGGCCTTTTC 3850  
Db -----AGGTGGCACTTTTC 1905  
Qy 3851 GGGAAATGTGCGGAAACCCCTATTGTTTATTTTCTAAATACATTTCAAAATGATGATC 3910  
Db GGGAAATGTGCGGAAACCCCTATTGTTTATTTTCTAAATACATTTCAAAATGATGATC 1965  
Qy 3911 CGCTCATGAGCAATTAACCTGATAAATGCTTCAATAATTAATGAAGAAGAGTATGA 3970  
Db CGCTCATGAGCAATTAACCTGATAAATGCTTCAATAATTAATGAAGAAGAGTATGA 2025  
Qy 3971 GTATTCAATTTTCGCTGCGCTTATTCCTTTTTCGGCAATTTTTCCTTCTCTGTTT 4030  
Db GTATTCAATTTTCGCTGCGCTTATTCCTTTTTCGGCAATTTTTCCTTCTCTGTTT 2085  
Qy 4031 TTGCTCACCCAGAAAACCTGCTGAAGATAAAGATGCTGAAGATCAGTTGGGTGCAAG 4090  
Db TTGCTCACCCAGAAAACCTGCTGAAGATAAAGATGCTGAAGATCAGTTGGGTGCAAG 2145  
Qy 4091 TGGGTTCATCGAATCGATCTCAACAGCGGTGAAGATCTCTGAGAGTTTTCGCCCGAAG 4150  
Db TGGGTTCATCGAATCGATCTCAACAGCGGTGAAGATCTCTGAGAGTTTTCGCCCGAAG 2205  
Qy 4151 AAGTTTTTCAATGATGAGCACTTTTAAAGTTCTGCTATGCGGCTATTTATCCGTA 4210  
Db AAGTTTTTCAATGATGAGCACTTTTAAAGTTCTGCTATGCGGCTATTTATCCGTA 2265  
Qy 4211 TTGACGCGCGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTC 4270  
Db TTGACGCGCGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTC 2325  
Qy 4271 AGTACTCACCAAGCAATCTTACGATGGCATGACAGTAAGAGAAATTAAGCA 4330

Db 2326 AGTACTCACCAAGCAATCTTACGATGGCATGACAGTAAGAGAAATTAAGCA 2385  
Qy 4331 GTGCTGCATTAACCATGAGTGATTAACACTGCGGCCAACTTACTTCTGACAAAGATCGAG 4390  
Db GTGCTGCATTAACCATGAGTGATTAACACTGCGGCCAACTTACTTCTGACAAAGATCGAG 2445  
Qy 4391 GACCGAAGGAGCTTAAACCGCTTTTTCGCAACAATGCGGGATCATGTAACCTGCGCTTATC 4450  
Db GACCGAAGGAGCTTAAACCGCTTTTTCGCAACAATGCGGGATCATGTAACCTGCGCTTATC 2505  
Qy 4451 GTTGGGAACCGGAGCTGAATGAAGCCATACAAACGACGAGCGTGACACCAAGATGCTG 4510  
Db GTTGGGAACCGGAGCTGAATGAAGCCATACAAACGACGAGCGTGACACCAAGATGCTG 2565  
Qy 4511 TAGCAATGCAACAAACGCTGCGCAAACTTATTAACCTGCGGAACTTACTTCTAGCTTCCC 4570  
Db TAGCAATGCAACAAACGCTGCGCAAACTTATTAACCTGCGGAACTTACTTCTAGCTTCCC 2625  
Qy 4571 GGCACAAATTAATAGACTGGATGGAGCGGATAAAGTTGACGACCACTTCTGCGCTCGG 4630  
Db GGCACAAATTAATAGACTGGATGGAGCGGATAAAGTTGACGACCACTTCTGCGCTCGG 2685  
Qy 4631 CCTTCCGCTGCGTGTATTTGCTGATTAATCTGAGCGCGTGAGCGTGGTCTCGCG 4690  
Db CCTTCCGCTGCGTGTATTTGCTGATTAATCTGAGCGCGTGAGCGTGGTCTCGCG 2745  
Qy 4691 GTATCATTTGAGCACTGGGCGCGAGATGTAAGCCCTCCCTGATCTGTAATCTACACGA 4750  
Db GTATCATTTGAGCACTGGGCGCGAGATGTAAGCCCTCCCTGATCTGTAATCTACACGA 2805  
Qy 4751 CGGGAGTCAAGCAACTATGATGAACAAATAGACAGATCGCTGAGATAGTGTCTCAC 4810  
Db CGGGAGTCAAGCAACTATGATGAACAAATAGACAGATCGCTGAGATAGTGTCTCAC 2865  
Qy 4811 TGATTAAGCAATGCTAACTCTCAGACCAAGTTTACTCATATATATCTTTAGATTTGATTA 4870  
Db TGATTAAGCAATGCTAACTCTCAGACCAAGTTTACTCATATATATCTTTAGATTTGATTA 2925  
Qy 4871 AACTTCAATTTTAAATTAAGGATCTAGTGAGATCTTTTGTGATAATCTCATGACCA 4930  
Db AACTTCAATTTTAAATTAAGGATCTAGTGAGATCTTTTGTGATAATCTCATGACCA 2985  
Qy 4931 AAATCCCTTAACTGAGTTTTCCTTCCACTGAGCGTCAGACCCCGTGAAGAAAGTCAAAG 4990  
Db AAATCCCTTAACTGAGTTTTCCTTCCACTGAGCGTCAGACCCCGTGAAGAAAGTCAAAG 3045  
Qy 4991 GATCTTCTTGAGATCTCTTTTTCGCGCTAATCTGCTGCTTGCACAAACAAACCAAC 5050  
Db GATCTTCTTGAGATCTCTTTTTCGCGCTAATCTGCTGCTTGCACAAACAAACCAAC 3105  
Qy 5051 CGCTTACCGCGGTGTTTTCGCGGATCAAGAGCTTACCACTCTTTTTCGGAAGGTAA 5110  
Db CGCTTACCGCGGTGTTTTCGCGGATCAAGAGCTTACCACTCTTTTTCGGAAGGTAA 3165  
Qy 5111 CTGCTTTCAGCAGAGCGAGATACCAATCTGCTTCTAGTGTAGCGGTAGTTCAGCC 5170  
Db CTGCTTTCAGCAGAGCGAGATACCAATCTGCTTCTAGTGTAGCGGTAGTTCAGCC 3225  
Qy 5171 ACCACTTCAAGAACTCTGTAGCAACCGCTTACATCTGCTTCTGCTAATCTGTTACAG 5230  
Db ACCACTTCAAGAACTCTGTAGCAACCGCTTACATCTGCTTCTGCTAATCTGTTACAG 3285  
Qy 5231 TGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5290  
Db TGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3345  
Qy 5291 CGGATTAAGCGCAGCGCTGAGCGGGGTTTCTGTCACACAGCCAGCTTGGAGC 5350  
Db CGGATTAAGCGCAGCGCTGAGCGGGGTTTCTGTCACACAGCCAGCTTGGAGC 3405  
Qy 5351 GAACGACCTTACACCGAATCTGAGATCTTACAGCGGTGAGCTTATGAGAAAGCGCCCTTC 5410

Db	3406	GAACGACCTTACACCAACTGAGATACCTTAAGCGTGAGCTATGAGAAAGCCGACGGTTC	3465
Qy	5411	CCGAAGGGAGAAAGCGGACAGGATATCCGGTAAGCGGAGGGTCGGAAACAGGAGAGCGCA	5470
Db	3466	CCGAAGGGAGAAAGCGGACAGGATATCCGGTAAGCGGAGGGTCGGAAACAGGAGAGCGCA	3525
Qy	5471	CGAGGAGCTTCCAGGGGGAAACGCTGGTATCTTTATAGTCCCTGTGGGGTTTCGCCACC	5530
Db	3526	CGAGGAGCTTCCAGGGGGAAACGCTGGTATCTTTATAGTCCCTGTGGGGTTTCGCCACC	3585
Qy	5531	TCTGACTTGAGCGTCGATTTTGTGTGATGCTGCTCAGGGGGCGGAGCCTATGGAAAAACG	5590
Db	3586	TCTGACTTGAGCGTCGATTTTGTGTGATGCTGCTCAGGGGGCGGAGCCTATGGAAAAACG	3645
Qy	5591	CCAGCAACGCGGCGCTTTTACGGTTCTCGGCTTTTGTGCGCTTTTGTCTCACATGTTCT	5650
Db	3646	CCAGCAACGCGGCGCTTTTACGGTTCTCGGCTTTTGTGCGCTTTTGTCTCACATGTTCT	3705
Qy	5651	TTCTCGGTTATCCCTGATTTCTGTGATTAACCGTATTTACCGCTTTTGTGAGTGAGCTGATA	5710
Db	3706	TTCTCGGTTATCCCTGATTTCTGTGATTAACCGTATTTACCGCTTTTGTGAGTGAGCTGATA	3765
Qy	5711	CCGCTCCCGCAGCGGACGACGAGCGGACGAGTCACTGAGCGGAGCGGAGCGGAGGAGC	5770
Db	3766	CCGCTCCCGCAGCGGACGACGAGCGGACGAGTCACTGAGCGGAGCGGAGCGGAGGAGC	3825
Qy	5771	GCCCAATACGCAACCGGCTCTCCCGCGCTGCGCGGCTGCGCGGCTTCAATTAATGCAG	5822
Db	3826	GCCCAATACGCAACCGGCTCTCCCGCGCTGCGCGGCTTCAATTAATGCAG	3877
RESULT 15			
US-09-991-209-15			
; Sequence 15, Application US/09991209			
; Publication No. US20030024009A1			
; GENERAL INFORMATION:			
; APPLICANT: Dunn-Coleman, Nigel			
; APPLICANT: Langdon, Timothy			
; APPLICANT: Morse, Philip			
; TITLE OF INVENTION: Manipulation of the Phenolic Acid			
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted			
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes			
; FILE REFERENCE: G6648-2			
; CURRENT APPLICATION NUMBER: US/09/991,209			
; CURRENT FILING DATE: 2002-07-02			
; PRIOR APPLICATION NUMBER: US 60/249,608			
; PRIOR FILING DATE: 2000-11-17			
; NUMBER OF SEQ ID NOS: 97			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 15			
; LENGTH: 5338			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: pTP10-1 vector			
US-09-991-209-15			
Query Match 47.5%; Score 2768; DB 10; Length 5338;			
Best Local Similarity 96.8%; Pred. No. 0;			
Matches 2895; Conservative 0; Mismatches 25; Indels 72; Gaps 4;			
Qy	2835	AAGCAGATCGTTCAACACATTTGGCAATAAGTTTCTTAAGATTGAATCCTGTTGCCGGTC	2894
Db	955	AAGCAGATCGTTCAACACATTTGGCAATAAGTTTCTTAAGATTGAATCCTGTTGCCGGTC	1014
Qy	2895	TTGGATGATATATCATATAATTTCTGTTGAATTAGCTTAAGCTATTAATTAACATGT	2954
Db	1015	TTGGATGATATATCATATAATTTCTGTTGAATTAGCTTAAGCTATTAATTAACATGT	1074
Qy	2955	AATGCAATGAGTATTTATGAGTGGGTTTTTATGATTAGAGTCCCGCAATTAACATTT	3014
Db	1075	AATGCAATGAGTATTTATGAGTGGGTTTTTATGATTAGAGTCCCGCAATTAACATTT	1134

Qy	3015	AATACGCGATAGAAAAAACAATAATAGCGCGCAAACTAGGATAAATTTATCGCGCGGGTGT	3074
Db	1135	AATACGCGATAGAAAAAACAATAATAGCGCGCAAACTAGGATAAATTTATCGCGCGGGTGT	1194
Qy	3075	CATCTATGTTTACTAGATCGACCTCGCAGGCATGGGATCCGCGCGCCGATCGGACGTCGGGC	3134
Db	1195	CATCTATGTTTACTAGATCGAATAGCTTCTAGAGCGCGCGGTGGAGC-----T	1241
Qy	3135	CCAAATTCGCCCTATAGTGAGTCGTAATTAC---AATTTCACTGGCCGCTGTTTTTACAACGTC	3191
Db	1242	CCAAATTCGCCCTATAGTGAGTCGTAATTAC---AATTTCACTGGCCGCTGTTTTTACAACGTC	1301
Qy	3192	GTGACTGGGAAAACCCCTGGCGTTTACCCAACTTAATCGCTTTTCAGCACAATCCCCCTTTCG	3251
Db	1302	GTGACTGGGAAAACCCCTGGCGTTTACCCAACTTAATCGCTTTTCAGCACAATCCCCCTTTCG	1361
Qy	3252	CCAGCTGGCGTAATAGCGAGAGAGCCCGCACCGATCGCCCTTCCCAACAGTTTGGCGAGCC	3311
Db	1362	CCAGCTGGCGTAATAGCGAGAGAGCCCGCACCGATCGCCCTTCCCAACAGTTTGGCGAGCC	1421
Qy	3312	TGAATGGCGAAT-GGACGCGCCCTGTAGCGCGCATTAAGCGCGCGGCTGTGGTGTGA	3370
Db	1422	TGAATGGCGAATGGACGCGCCCTGTAGCGCGCATTAAGCGCGCGGCTGTGGTGTGA	1481
Qy	3371	CGCGAGCGTGACCGCTACACTTGCAGCGCCCTAGCGCCCGCTCCTTTTGGCTTTCTTCC	3430
Db	1482	CGCGAGCGTGACCGCTACACTTGCAGCGCCCTAGCGCCCGCTCCTTTTGGCTTTCTTCC	1541
Qy	3431	CTTCTTTTCTGCCACGTTTCCCGGCTTTTCCCGTCAAGCTCTAAATCGGGGGCTCCCTT	3490
Db	1542	CTTCTTTTCTGCCACGTTTCCCGGCTTTTCCCGTCAAGCTCTAAATCGGGGGCTCCCTT	1601
Qy	3491	TAGGGTTCCGATTTTAGAGCTTTTACGGCACCTCGACCGCAAAAAAACTTGATTTGGGTGATG	3550
Db	1602	TAGGGTTCCGATTTTAGAGCTTTTACGGCACCTCGACCGCAAAAAAACTTGATTTGGGTGATG	1661
Qy	3551	GTTCACTAGTGGGCGAATCGCCCTGTAGAGCGGTTTTTTCGCCCTTTTGGAGTTGGAGTCCA	3610
Db	1662	GTTCACTAGTGGGCGAATCGCCCTGTAGAGCGGTTTTTTCGCCCTTTTGGAGTTGGAGTCCA	1721
Qy	3611	CGTTCTTTTAAATAGTGGACTCTTGTTCGAACTGGAACAACTCAACCTATCTCGGTCT	3670
Db	1722	CGTTCTTTTAAATAGTGGACTCTTGTTCGAACTGGAACAACTCAACCTATCTCGGTCT	1781
Qy	3671	ATTCCTTTTGAATTAAGGGGATTTTGGCGATTTTCGGCTATTTGGTTAAAAAATAGCTGA	3730
Db	1782	ATTCCTTTTGAATTAAGGGGATTTTGGCGATTTTCGGCTATTTGGTTAAAAAATAGCTGA	1841
Qy	3731	TTTAAACAAATTTTAAACCGGAATTTTAAACAAATAATTAACGTTTTCGCTGATG	3790
Db	1842	TTTAAACAAATTTTAAACCGGAATTTTAAACAAATAATTAACGTTTTCGCTGATG	1892
Qy	3791	CGGTAATTTTCTCCTTACGCACTCTGCGGTATTTTACACCGCATACAGTGGCACTTTTC	3850
Db	1893	CGGTAATTTTCTCCTTACGCACTCTGCGGTATTTTACACCGCATACAGTGGCACTTTTC	1906
Qy	3851	GGGGAAATGTGCGGGAACCCCTAATTTTCTTAAATACATTTCAAAATATGATATC	3910
Db	1907	GGGGAAATGTGCGGGAACCCCTAATTTTCTTAAATACATTTCAAAATATGATATC	1966
Qy	3911	CGCTCATGAGCAATAACCCCTGATAAATGCTTCAATAATATGAAAAAGGAGATGTA	3970
Db	1967	CGCTCATGAGCAATAACCCCTGATAAATGCTTCAATAATATGAAAAAGGAGATGTA	2026
Qy	3971	GTATTTCAACATTTCCGTTGTCCTTATTTCCCTTTTTCGCGCATTTTTCCTGTTT	4030
Db	2027	GTATTTCAACATTTCCGTTGTCCTTATTTCCCTTTTTCGCGCATTTTTCCTGTTT	2086
Qy	4031	TTGCTCAACCAAGAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGACGAG	4090
Db	2087	TTGCTCAACCAAGAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGACGAG	2146
Qy	4091	TGGGTTACTCGAAGTGAATCTCAACAGCGGTAGATCCTTTGAGAGTTTTCGCGCGGAG	4150



|||||  
2147 TGGGTTACATCGAATCGGATCTCAACAGCGGTAAAGTCTTGAAGTTTTCGCCCGAAG 2206 Db  
|||||  
4151 AACGTTTTCCAAATGATGAGCACTTTTAAAGTTCTGTATATGCGCGGTATTAATCCGGTA 4210 Qy  
|||||  
2207 AACGTTTTCCAAATGATGAGCACTTTTAAAGTTCTGTATATGCGCGGTATTAATCCGGTA 2266 Db  
|||||  
4211 TTGACCGCGGCAAGAGCACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTG 4270 Qy  
|||||  
2267 TTGACCGCGGCAAGAGCACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTG 2326 Db  
|||||  
4271 AGTACTCACCAGTACACAGAAAAGCATCTTACCGATGCGATGACAGTAAGAGAAATTAAGCA 4330 Qy  
|||||  
2327 AGTACTCACCAGTACACAGAAAAGCATCTTACCGATGCGATGACAGTAAGAGAAATTAAGCA 2386 Db  
|||||  
4331 GTGCTGCCATTAACCATGAGTGATAACACTCGCGGCCAACTTTACTTCTGACAAACGATCGGAG 4390 Qy  
|||||  
2387 GTGCTGCCATTAACCATGAGTGATAACAACCTCGCGGCCAACTTTACTTCTGACAAACGATCGGAG 2446 Db  
|||||  
4391 GACCGAAGGAGCTAAACCGCTTTTGGCAACAATGCGGGATCATGTAACCTCGCCTTGATC 4450 Qy  
|||||  
2447 GACCGAAGGAGCTAAACCGCTTTTGGCAACAATGCGGGATCATGTAACCTCGCCTTGATC 2506 Db  
|||||  
4451 GTTGGGAAACCGGAGCTGAATGAAGCCATACCAACAGACAGCGTGAACACCAACGATGCGCTG 4510 Qy  
|||||  
2507 GTTGGGAAACCGGAGCTGAATGAAGCCATACCAACAGACAGCGTGAACACCAACGATGCGCTG 2566 Db  
|||||  
4511 TAGCAATGGCAACAACGTTGCGCAAACTATTAACTGCGCAACTACTTTACTCTAGCTTCCC 4570 Qy  
|||||  
2567 TAGCAATGGCAACAACGTTGCGCAAACTATTAACTGCGCAACTACTTTACTCTAGCTTCCC 2626 Db  
|||||  
4571 GGCACAAATTAATAGACTGGATGGAGCGGATAAAGTTGACAGACCACTTCTCGGCTCGG 4630 Qy  
|||||  
2627 GGCACAAATTAATAGACTGGATGGAGCGGATAAAGTTGACAGACCACTTCTCGGCTCGG 2686 Db  
|||||  
4631 CCCTTCGCGTGGCTGTTTATTGCTGATAAATCTGGAGCGGTGAGCGTGGGTCTCGCG 4690 Qy  
|||||  
2687 CCCTTCGCGTGGCTGTTTATTGCTGATAAATCTGGAGCCGTGAGCGTGGGTCTCGCG 2746 Db  
|||||  
4691 GTATCATTTGACGACCTGGGCGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGA 4750 Qy  
|||||  
2747 GTATCATTTGACGACCTGGGCGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGA 2806 Db  
|||||  
4751 CGGGAGTCAGGCAACTATCGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCAC 4810 Qy  
|||||  
2807 CGGGAGTCAGGCAACTATCGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCAC 2866 Db  
|||||  
4811 TGATTAAAGCATTCGTAAGTCTGAGCAACCAAGTTTACTCATATATACTTTAGATTGATTAA 4870 Qy  
|||||  
2867 TGATTAAAGCATTCGTAAGTCTGAGCAACCAAGTTTACTCATATATACTTTAGATTGATTAA 2926 Db  
|||||  
4871 AACTTCATTTTTTAAATGAGATCTAGTGAAGATCCTTTTGTGATPAATCTCATGACCA 4930 Qy  
|||||  
2927 AACTTCATTTTTTAAATGAGATCTAGTGAAGATCCTTTTGTGATPAATCTCATGACCA 2986 Db  
|||||  
4931 AAATCCCTTAAAGTGGTTTTGTTTCACTGAGCGTCAGACCCCGTAGAAAAGATCAAG 4990 Qy  
|||||  
2987 AAATCCCTTAAAGTGGTTTTGTTTCACTGAGCGTCAGACCCCGTAGAAAAGATCAAG 3046 Db  
|||||  
4991 GATCTTCTTCAGATCCTTTTCTGCGGTAACTGCTGCTGCAACCAAAAACCAAC 5050 Qy  
|||||  
3047 GATCTTCTTCAGATCCTTTTCTGCGGTAACTGCTGCTGCAACCAAAAACCAAC 3106 Db  
|||||  
5051 CGCTACAGCGGTGGTTTTGTTTCCCGGATCAAGAGCTTACCAACTTTTTTCCGAAGGTAA 5110 Qy  
|||||  
3107 CGCTACAGCGGTGGTTTTGTTTCCCGGATCAAGAGCTTACCAACTTTTTTCCGAAGGTAA 3166 Db  
|||||  
5111 CTGGCTTCAGCAGAGCGCAGATACCAAAATCTGCTTCTAGTGTAGCGGTAGTAGGCC 5170 Qy  
|||||  
3167 CTGGCTTCAGCAGAGCGCAGATACCAAAATCTGCTTCTAGTGTAGCGGTAGTAGGCC 3226 Db  
|||||  
5171 ACCACTTCAGAACTCTGTAGCAGCGCTACATACCTCGCTCTGCTAATCTCTGTACCAG 5230 Qy  
|||||

3227 ACCACTTCAGAACTCTGTAGCAGCGCTACATACCTCGCTCTGCTAATCTCTGTACCAG 3286 Db  
|||||  
5231 TGGCTGCTGCACTGCGGATAAAGTCTGTCTTTACCGGGTTGGACTCAAGAGCATAGTTAC 5290 Qy  
|||||  
3287 TGGCTGCTGCACTGCGGATAAAGTCTGTCTTTACCGGGTTGGACTCAAGAGCATAGTTAC 3346 Db  
|||||  
5291 CGGATAAGCGCGCAGCGGTGCGGCTGAAACGGGGGTTCTGTCACACAGCCCGCAGCTTGGAGC 5350 Qy  
|||||  
3347 CGGATAAGCGCGCAGCGGTGCGGCTGAAACGGGGGTTCTGTCACACAGCCCGCAGCTTGGAGC 3406 Db  
|||||  
5351 GAAACGACCTACACCGAACTGAGATACCTACAGCGTGTAGCTTATGAGAAAGGCCACGCTTC 5410 Qy  
|||||  
3407 GAAACGACCTACACCGAACTGAGATACCTACAGCGTGTAGCTTATGAGAAAGGCCACGCTTC 3466 Db  
|||||  
5411 CCAGAGGAGAGAAAGGCGGACAGGTATCCCGTAAAGCGGTCGGAACAGAGAGCGCA 5470 Qy  
|||||  
3467 CCAGAGGAGAGAAAGGCGGACAGGTATCCCGTAAAGCGGTCGGAACAGAGAGCGCA 3526 Db  
|||||  
5471 CGAGGAGAGCTTCCAGGGGGGAAACGCTGTATCTTTATAGTCTCTGTCGGGTTTCGCCACC 5530 Qy  
|||||  
3527 CGAGGAGAGCTTCCAGGGGGGAAACGCTGTATCTTTATAGTCTCTGTCGGGTTTCGCCACC 3586 Db  
|||||  
5531 TCTGACTTCGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAAAAAAG 5590 Qy  
|||||  
3587 TCTGACTTCGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAAAAAAG 3646 Db  
|||||  
5591 CGAGCAACCGCGGCTTTTAAOGGTTCTCGGCTTTTGTGCGCTTTTGTCTCAGATGTTCT 5650 Qy  
|||||  
3647 CGAGCAACCGCGGCTTTTAAOGGTTCTCGGCTTTTGTGCGCTTTTGTCTCAGATGTTCT 3706 Db  
|||||  
5651 TTCCTCGGTTATCCCTGATTTCTGTGATAAACCGTATTACCGCTTTTGTGAGTGTGATGATA 5710 Qy  
|||||  
3707 TTCCTCGGTTATCCCTGATTTCTGTGATAAACCGTATTACCGCTTTTGTGAGTGTGATGATA 3766 Db  
|||||  
5711 CCCTCGCGCAGCCGAAACGACCGAGCGCAGCGAGTCAGTGAGCGAGAAAGCGGAAGAGC 5770 Qy  
|||||  
3767 CCCTCGCGCAGCCGAAACGACCGAGCGCAGCGAGTCAGTGAGCGAGAAAGCGGAAGAGC 3826 Db  
|||||  
5771 GCCCAATACGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGCGAG 5822 Qy  
|||||  
3827 GCCCAATACGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGCGAG 3878 Db  
|||||

Search completed: June 20, 2004, 15:18:02  
Job time : 2242 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 10:21:08 ; Search time 21590 Seconds  
(without alignments)  
11687.964 Million cell updates/sec

Title: US-09-924-197-1

Perfect score: 5822

Sequence: 1 ctggacagacagggttccg.....tggccgattcattaatgag 5822

Scoring table: IDENTITY\_NUC

Gapop 10\_0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_bt.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_hcg\_hum.\*

31: em\_hcg\_inv.\*

32: em\_hcg\_other.\*

33: em\_hcg\_mus.\*

34: em\_hcg\_pln.\*

35: em\_hcg\_rtd.\*

36: em\_hcg\_mam.\*

37: em\_hcg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2892	49.7	4486	12	EVCOR116N	229589 Expression
2	2760.8	47.4	4412	12	EVCOR112N	229587 Expression
3	2755.6	47.3	5534	6	AR037157	AR037157 Sequence
4	2755.6	47.3	5534	6	AR112043	AR112043 Sequence
5	2705.8	46.5	4350	12	AF173954	AF173954 Cloning v
6	2705.8	46.5	5558	12	AF173955	AF173955 Cloning v
7	2686	46.1	4229	6	E63777	E63777 Method for
8	2683.4	46.1	3448	6	AX590988	AX590988 Sequence
9	2683.4	46.1	3448	6	AX591141	AX591141 Sequence
10	2683.4	46.1	3448	6	AX717564	AX717564 Sequence
11	2683.4	46.1	3551	12	AY122058	AY122058 Expression
12	2683.4	46.1	3711	12	AF525778	AF525778 Expression
13	2677	46.0	3033	12	CVGEM7LICF	U25268 Ligatation-in
14	2677	46.0	3033	12	CVGEM7LICR	U25268 Ligatation-in
15	2677	46.0	7823	12	AF041426	AF041426 Cloning v
16	2677	46.0	8658	12	AF338824	AF338824 Cloning v
17	2677	46.0	8696	12	AF379854	AF379854 Cloning v
18	2675.4	46.0	3018	6	A98767	A98767 Sequence 4
19	2671.6	45.9	3877	12	AB038599	AB038599 Cloning v
20	2671.6	45.9	6320	12	AB038600	AB038600 Cloning v
21	2668.6	45.8	4514	6	AX781452	AX781452 Sequence
22	2668.6	45.8	4514	6	AX816961	AX816961 Sequence
23	2658.4	45.7	9359	6	AX384394	AX384394 Sequence
24	2658.4	45.7	9359	6	AX473364	AX473364 Sequence
25	2654.6	45.6	2943	12	AF092940	AF092940 Cloning v
26	2654.6	45.6	3012	12	AF092546	AF092546 Cloning v
27	2641.4	45.4	2997	12	CVGEM7ZFP	X65310 Cloning vec
28	2641.4	45.4	3000	12	CVGEM5ZRP	X65308 Cloning vec
29	2641.4	45.4	3404	6	AX771236	AX771236 Sequence
30	2630.4	45.2	3485	6	AR199035	AR199035 Sequence
31	2553	43.9	6824	6	A25909	A25909 Synthetic y
32	2552	43.8	10138	12	AF187951	AF187951 Activatio
33	2552	43.8	10450	12	AF218466	AF218466 Activatio
34	2551.8	43.8	37808	6	AX001082	AX001082 Sequence
35	2550	43.8	4883	6	AR220205	AR220205 Sequence
36	2550	43.8	4883	6	AX358365	AX358365 Sequence
37	2550	43.8	11038	12	AY196826	AY196826 Piggybac
38	2549	43.8	2958	6	AX247549	AX247549 Sequence
39	2549	43.8	2958	6	AX247550	AX247550 Sequence
40	2549	43.8	2958	6	AX247551	AX247551 Sequence
41	2549	43.8	2967	12	U02449	U02449 Cloning vec
42	2547.4	43.8	4133	12	U01668	U01668 Phagemid cl
43	2542.6	43.7	2959	12	AF118920	AF118920 Cloning v
44	2538.2	43.6	10597	6	AX006825	AX006825 Sequence
45	2538.2	43.6	10597	6	AX417673	AX417673 Sequence

#### ALIGNMENTS

RESULT 1  
EVCOR116N  
LOCUS EVCOR116N 4486 bp DNA linear SYN 24-MAR-1994  
DEFINITION Expression vector pCOR116N (modified from pCOR116 in [3]).  
ACCESSION 229589  
VERSION 229589.1 GI:452348  
KEYWORDS ampicillin resistance; beta-lactamase; colE1 origin; expression  
vector; nos terminator; phage fl region; rice actin1 promoter.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 49 to 304)  
AUTHORS Depicker,A., Stachel,S., Dhaese,P., Zambryski,P. and Goodman,H.M.  
TITLE Napaline synthase: transcript and DNA sequence  
JOURNAL J. Gen. Appl. Microbiol. 1, 561-573 (1982)



2 McElroy, D., Blowers, A.D., Jones, B. and Wu, R.  
 TITLE Construction of expression vectors based on the rice actin 1 (Act1)  
 5' region for use in monocot transformation  
 JOURNAL Mol. Gen. Genet. 231 (1), 150-160 (1991)  
 MEDLINE 92092956  
 PUBMED 1753941  
 REMARK (sites)  
 3 (bases 1 to 4486)  
 AUTHORS Liu, L., Dasgupta, I., Davies, J. and Hull, R.  
 TITLE Modified expression vectors for monocot transformation toward virus  
 resistance  
 JOURNAL Unpublished  
 REFERENCE 4 (bases 1 to 4486)  
 AUTHORS Liu, L.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-JAN-1994) Liu L., John Innes Institute, Virus  
 Research, Colney Lane, Norwich, United Kingdom, NR4 7UH  
 FEATURES  
 source  
 1. 4486  
 /organism="synthetic construct"  
 /mol\_type="other DNA"  
 /db\_xref="taxon:32630"  
 misc\_feature  
 1. 52  
 /note="multiple cloning sites"  
 misc\_feature  
 3. 5  
 /note="start codon for gene expression"  
 terminator  
 49. 304  
 /note="nos terminator"  
 /citation=[1]  
 misc\_feature  
 511. 967  
 /note="phage f1 region"  
 CDS  
 1153. 2013  
 /note="unnamed protein product; beta-lactamase coding  
 region"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAA82706.1"  
 /db\_xref="GI:452349"  
 /db\_xref="REMYEMBL:CAA82706"  
 /translation="MSIQHFRVALIPFPAAFLPVFAHPETLVKVKDAEDQLGARVGY  
 IELDLSKRILESFPEERPFMSDFKVLICHLLSRIDAGQQLGRPRYSQNDLVE  
 YSPVTERHLDGMTVRELCSAAITMSDNTAANLLLTIGGPKRLTAFIHNMGDHYRL  
 DRWPELNEALPNDREDYTPVAMPTTLKLLTGLLETLASRQQLIDHMEADKVGPL  
 LRSALPAGWFTADSKGAGERGRGIIAALGPDGKPSRIWIYITGSGQATMDERNQIA  
 EIGASLIKHW"  
 misc\_feature  
 2014. 2954  
 /note="ColEI origin"  
 promoter  
 3256. 4486  
 /note="rice actin 1 5' region, (act1 promoter)"  
 /citation=[2]  
 ORIGIN  
 Query Match 49.7%; Score 2892; DB 12; Length 4486;  
 Best Local Similarity 98.5%; Pred. No. 0;  
 Matches 2940; Conservative 0; Mismatches 20; Indels 26; Gaps 1;  
 QY 2837 GCAGATCGTTCAAAATTTGGCAATAAGTTTCTTAAGATTGAATCCTGTTGCCGCTCTT 2896  
 DB 49 GCAGATCGTTCAAAATTTGGCAATAAGTTTCTTAAGATTGAATCCTGTTGCCGCTCTT 108  
 QY 2897 GCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTATAATTAACATGTAA 2956  
 DB 109 GCGATGATTATCATATAATTTCTGTTGAATTAGCTTAAGCATGTATAATTAACATGTAA 168  
 QY 2957 TGCATGAGCTTATTTATGAGATGGGTTTTATGATTAGATAGTCCCGCAATTATACATTTAA 3016  
 DB 169 TGCATGAGCTTATTTATGAGATGGGTTTTATGATTAGATAGTCCCGCAATTATACATTTAA 228  
 QY 3017 TACGCGATAGAAACAAATATAGCGCGCAACTAGGATAAATTATCGCGCGGTGTCA 3076  
 DB 229 TACGCGATAGAAACAAATATAGCGCGCAACTAGGATAAATTATCGCGCGGTGTCA 288

3077 TCTATGTTACTAGATCGACCTGCGAGCATGGGATCCGCGCGCATGCGATGCGAGTCTCGGGCCC 3136  
 DB TCTATGTTACTAGAT-----CTCTAGAGTGGAGCGGCCC 322  
 QY 3137 AATTGCGCCCTATAGTAGTCTGTATTAATTCATCTGCGCGCTGCTTTTACAAGCTGTGAC 3196  
 DB AATTGCGCCCTATAGTAGTCTGTATTAATTCATCTGCGCGCTGCTTTTACAAGCTGTGAC 382  
 QY 3197 TGGGAAAAACCTGCGCTTACCCAACTTAATCGCTTTCAGCAGCATCCCTCTTTCGCCAGC 3256  
 DB TGGGAAAAACCTGCGCTTACCCAACTTAATCGCTTTCAGCAGCATCCCTCTTTCGCCAGC 442  
 QY 3257 TGGGTAATAGCGAAGAGCGCCGACCGATCGCTTCCCAACAGTTCGCGAGCTGAAT 3316  
 DB TGGGTAATAGCGAAGAGCGCCGACCGATCGCTTCCCAACAGTTCGCGAGCTGAAT 502  
 QY 3317 GCGAATAGCAGCGCCCTGTAGCGCGCATTAAGGCGCGGGTGTGTGTTAGCGCA 3376  
 DB GCGAATAGCAGCGCCCTGTAGCGCGCATTAAGGCGCGGGTGTGTGTTAGCGCA 562  
 QY 3377 GCGTGACCGCTACACTTGGCAGCGCCCTAGCGCGCCCTCTTTCGCTTCTTCCCTTCT 3436  
 DB GCGTGACCGCTACACTTGGCAGCGCCCTAGCGCGCCCTCTTTCGCTTCTTCCCTTCT 622  
 QY 3437 TTCTGCGCACGTTGCGCGGCTTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTTTAGGCT 3496  
 DB TTCTGCGCACGTTGCGCGGCTTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTTTAGGCT 682  
 QY 3497 TCCGATTTAGAGCTTTAGCGCACCTCGACCGCAAAACTTGAATTTGGGTGATGTTTAC 3556  
 DB TCCGATTTAGAGCTTTAGCGCACCTCGACCGCAAAACTTGAATTTGGGTGATGTTTAC 742  
 QY 3557 GTAGTGGGCCATCGCCCTGATAGACGGTTTTCGCGCTTTCGACGTTCGAGTCCAGCTTCT 3616  
 DB GTAGTGGGCCATCGCCCTGATAGACGGTTTTCGCGCTTTCGACGTTCGAGTCCAGCTTCT 802  
 QY 3617 TTAATAGTGGACTCTGTGTTCCAACTGGAACAACTCAACCTATCTCGCTCTATCTT 3676  
 DB TTAATAGTGGACTCTGTGTTCCAACTGGAACAACTCAACCTATCTCGCTCTATCTT 862  
 QY 3677 TTGATTTAATAGGATTTTTCGCGCTATTTGGTTAAAAATAGCTGATTAAAC 3736  
 DB TTGATTTAATAGGATTTTTCGCGCTATTTGGTTAAAAATAGCTGATTAAAC 922  
 QY 3737 AAATATTTAAGCGGAATTTTAAACAAATATTAACGTTTCAAAATTCGCTGATGCGGTAT 3796  
 DB AAATATTTAAGCGGAATTTTAAACAAATATTAACGTTTCAAAATTCGCTGATGCGGTAT 982  
 QY 3797 TTTCTCCTTACGCATCTGTGCGGTATTTCAACCGCATACAGTGGCAGCTTTTGGGGA 3856  
 DB TTTCTCCTTACGCATCTGTGCGGTATTTCAACCGCATACAGTGGCAGCTTTTGGGGA 3042  
 QY 3857 ATGTGCGGGAACCCCTATTTGTTTATTTTCTAAATACATTTCAATATGATTCGCTCA 3916  
 DB ATGTGCGGGAACCCCTATTTGTTTATTTTCTAAATACATTTCAATATGATTCGCTCA 1102  
 QY 3917 TGAGCAATAACCCCTGATAAATGCTTCAATATTTGAAAAAGAGAGTATGAGTATTC 3976  
 DB TGAGCAATAACCCCTGATAAATGCTTCAATATTTGAAAAAGAGAGTATGAGTATTC 1162  
 QY 3977 AACATTTCCGTGTGCGCCCTTATTCCTTTTTCGCGCATTTTTCCTTCTCTTCTTTC 4036  
 DB AACATTTCCGTGTGCGCCCTTATTCCTTTTTCGCGCATTTTTCCTTCTCTTCTTTC 1222  
 QY 4037 ACCCAGAAACGCTGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGACGAGTGGGTT 4096  
 DB ACCCAGAAACGCTGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGACGAGTGGGTT 1282  
 QY 4097 ACATCGAATCGGATCTCAACGAGGTAAAGATCTTTGAGAGTTTTCGCCCCGAAGACGTT 4156  
 DB ACATCGAATCGGATCTCAACGAGGTAAAGATCTTTGAGAGTTTTCGCCCCGAAGACGTT 1342  
 QY 4157 TTCCATATGATGACACTTTTAAAGTTCTGCTATGTGGCGCGGTATATATCCCGTATTGACG 4216

1343 TTCAATGATGAGCACTTTTAAAGTTCTGCTATGTATACACTATTTATCCGTTATTGACG 1402  
4217 CCGGGCAAGAGCAACTCGGTCGCGCATACACTATTTCTCAGAAATGACTTGGTTGAGTACT 4276  
1403 CCGGGCAAGAGCAACTCGGTCGCGCGCGGTATTTCTCAGAAATGACTTGGTTGAGTACT 1462  
4277 CACAGTCAAGAAAAGCACTTACGATGGCATGACAGTAAGAGAAATTAATGAGTCTG 4336  
1463 CACAGTCAAGAAAAGCACTTACGATGGCATGACAGTAAGAGAAATTAATGAGTCTG 1522  
4337 CCATAACCATGAGTAAACACTCGCGCAACTTACTTCTGACAAACCATCGAGGACCGA 4396  
1523 CCATAACCATGAGTAAACACTCGCGCAACTTACTTCTGACAAACCATCGAGGACCGA 1582  
4397 AGGAGCTTAACCGCTTTTTCGCAAACTAGGGGATCATGTAACTCGCTTGTATGCTGGG 4456  
1583 AGGAGCTTAACCGCTTTTTCGCAAACTAGGGGATCATGTAACTCGCTTGTATGCTGGG 1642  
4457 AACCGGAGCTGAATGAAGCCATACCAACGAGCGGTGACACCAACGATGCTGTAGCAA 4516  
1643 AACCGGAGCTGAATGAAGCCATACCAACGAGCGGTGACACCAACGATGCTGTAGCAA 1702  
4517 TGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTTACTTCTAGCTTTCGCGCAAC 4576  
1703 TGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTTACTTCTAGCTTTCGCGCAAC 1762  
4577 AATTAAATAGACTGATGAGCGGATAAAGTTGACGACCACTTCTCGCTCGGCGCTTC 4636  
1763 AATTAAATAGACTGATGAGCGGATAAAGTTGACGACCACTTCTCGCTCGGCGCTTC 1822  
4637 CGGCTGCTGTTTATGCTGATAAATCTGAGCGGTGAGCGGTGCTGCGGTATCA 4696  
1823 CGGCTGCTGTTTATGCTGATAAATCTGAGCGGTGAGCGGTGCTGCGGTATCA 1882  
4697 TTGAGCACTGCGGCGAGTGAAGCCCTCCCGTATCGTATGTTATCTACACGAGGGA 4756  
1883 TTGAGCACTGCGGCGAGTGAAGCCCTCCCGTATCGTATGTTATCTACACGAGGGA 1942  
4757 GTGAGCACTATGATGAACGAATAGACAGATCGCTGAGATAGGTGCTCTCACTGATTA 4816  
1943 GTGAGCACTATGATGAACGAATAGACAGATCGCTGAGATAGGTGCTCTCACTGATTA 2002  
4817 AGCATGCTGAATCTGTCAGACAAAGTTTATCTCATATATATCTTTAGATGATTTAAACCTC 4876  
2003 AGCATGCTGAATCTGTCAGACAAAGTTTATCTCATATATATCTTTAGATGATTTAAACCTC 2062  
4877 ATTTTAAATTTAAAGGATCTAGTGAAGATCTTTTGTATCTCATGACCAAAATCC 4936  
2063 ATTTTAAATTTAAAGGATCTAGTGAAGATCTTTTGTATCTCATGACCAAAATCC 2122  
4937 CTTAACGCTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAGGATCTT 4996  
2123 CTTAACGCTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAGGATCTT 2182  
4997 CTGAGATCTCTTTTTCGCGGTAAATCTGCTCTGCAAAACCAAAACCAACCGCTAC 5056  
2183 CTGAGATCTCTTTTTCGCGGTAAATCTGCTCTGCAAAACCAAAACCAACCGCTAC 2242  
5057 CAGCGGTGTTTTCGCGGTAAATCTGCTCTGCAAAACCAAAACCAACCGCTAC 5116  
2243 CAGCGGTGTTTTCGCGGTAAATCTGCTCTGCAAAACCAAAACCAACCGCTAC 2302  
5117 TCAGCAGCGCGCATACCAATCTGCTCTGCTAGTGTAGCGGTAGTGGCCACCACT 5176  
2303 TCAGCAGCGCGCATACCAATCTGCTCTGCTAGTGTAGCGGTAGTGGCCACCACT 2362  
5177 TCAAGAACTCTGTAGCAGCGCTACATACCTGCTCTGCTAACTCTGTTTACAGTGGCTG 5236  
2363 TCAAGAACTCTGTAGCAGCGCTACATACCTGCTCTGCTAACTCTGTTTACAGTGGCTG 2422  
5237 CTGCGAGTGGGATAGTCTGCTTACCGGTTGGATCTCAAGACCATAGTTTACCGGATA 5296

2423 CTGCCAGTGGCGATAAGTCTGTTTACCGGTTGGACTCAAGACGATAGTTTACCGGATA 2482  
5297 AGGCGCAGCGTTCGGCTGAAACCGGGGTTGCTGCAACACGCCAGCTTGGAGCGAAACA 5356  
2483 AGGCGCAGCGTTCGGCTGAAACCGGGGTTGCTGCAACACGCCAGCTTGGAGCGAAACA 2542  
5357 CCTACACCGCAACTGAGATACCTACAGCGTGTAGCTATGAGAAAGCGCCACGCTTCCCGAAG 5416  
2543 CCTACACCGCAACTGAGATACCTACAGCGTGTAGCTATGAGAAAGCGCCACGCTTCCCGAAG 2602  
5417 GGAGAAAGCGCGACAGCTATCCGTTAAGCGCAGCGTTCGAAACAGGAGAGCGCACGAGGG 5476  
2603 GGAGAAAGCGCGACAGCTATCCGTTAAGCGCAGCGTTCGAAACAGGAGAGCGCACGAGGG 2662  
5477 AGCTTCAGGGGGAACCGCTGCTATCTTATAGTCTCTGTCGGGTTTCGCCACCTCTGAC 5536  
2663 AGCTTCAGGGGGAACCGCTGCTATCTTATAGTCTCTGTCGGGTTTCGCCACCTCTGAC 2722  
5537 TTGAGCGTCCATTTTCTGATGCTGCTGCGGGGCGGAGCTATGGAACACGCCACGA 5596  
2723 TTGAGCGTCCATTTTCTGATGCTGCTGCGGGGCGGAGCTATGGAACACGCCACGA 2782  
5597 AGCGGCTCTTTTACCGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5656  
2783 AGCGGCTCTTTTACCGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2842  
5657 CGTTATCCCTGATCTTCTGATGAACCGTATTTACCGCTTTTGAAGTGAAGTGAAGTGAAG 5716  
2843 CGTTATCCCTGATCTTCTGATGAACCGTATTTACCGCTTTTGAAGTGAAGTGAAGTGAAG 2902  
5717 GCGGACCGCAACGACCGGCGAGCGAGTCAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAG 5776  
2903 GCGGACCGCAACGACCGGCGAGCGAGTCAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAG 2962  
5777 TAGCAAAACCGCTCTCCCGCGGTTGGCGGATTTCAATTAATGCAG 5822  
2963 TAGCAAAACCGCTCTCCCGCGGTTGGCGGATTTCAATTAATGCAG 3008

## RESULT 2

EVCOR112N  
LOCUS 4412 bp DNA linear SYN 18-MAR-1996  
DEFINITION Expression vector pCOR112N (modified from pCOR112 in [2]).  
ACCESSION Z29587  
VERSION Z29587.1 GI:452344  
KEYWORDS ampicillin resistance; beta-lactamase; colEI origin; expression vector; nos terminator; phage f1 region; rice actin1 promoter.  
SOURCE synthetic construct  
ORGANISM artificial sequence.  
REFERENCE 1 (bases 27 to 282)  
AUTHORS Depicker,A., Stachel,S., Dhaese,P., Zambryski,P. and Goodman,H.M.  
TITLE Nopaline synthase: transcript mapping and DNA sequence  
JOURNAL J. Mol. Appl. Genet. 1 (6), 561-573 (1982)  
MEDLINE 83110551  
PUBMED 7153689  
REFERENCE 2  
AUTHORS McElroy,D., Blowers,A.D., Jenes,B. and Wu,R.  
TITLE Construction of expression vectors based on the rice actin 1 (Act1) 5' region for use in monocot transformation  
JOURNAL Mol. Gen. Genet. 231 (1), 150-160 (1991)  
MEDLINE 92092956  
PUBMED 1753941  
REFERENCE 3  
AUTHORS Liu,L., Dasgupta,I., Davies,J.W. and Hull,R.  
TITLE Modified vectors for monocot transformation toward virus resistance  
JOURNAL Unpublished  
PUBMED 4 (bases 1 to 4412)  
AUTHORS Liu,L.  
TITLE Direct Submission  
JOURNAL Submitted (26-JAN-1994) Liu L., John Innes Institute, Virus Research, Colney Lane, Norwich, United Kingdom, NR4 7UH

## FEATURES

Source Location/Qualifiers  
1..4412  
/organism="synthetic construct"  
/mol\_type="other DNA"  
/db\_xref="taxon:32630"  
misc\_feature 1..30  
misc\_feature 3..5  
terminator /note="multiple cloning sites"  
27..282  
/note="start codon for gene expression"  
/note="nos terminator"  
/citation=[1]  
misc\_feature 506..962  
CDS /note="phage f1 region"  
1093..1953  
/codon\_start=1  
/transl\_table=11  
/product="beta-lactamase"  
/protein\_id="CAA82704.1"  
/db\_xref="GI:452345"  
/db\_xref="RBMTRMBL:CAA82704"  
/translation="MSIQHFRVALIPFAAPCLPVFAHPETLVKVKDAEDQLGARVGY  
IELDLNSGKILSPREEPRMWMSTEKVLICGLAVLSRIDAGQEQQLGRRIHYQNDLVE  
YSPVTEKHDTGIVTRELCSAALTMSDNTAANLLGLTTGPKELTAFLRWMDGVKRL  
DRWPELNEAIPIVDERDITMPVAMATTLKLLTGLLTLASRQQLIDWMEADKVAQL  
LRSLAPAGHFIADKSGAGSRGRIIAALGPDKPSRIIVIVITGSGQATWDERNRQIA  
EIGASLIKHW"  
1954..2894  
misc\_feature /note="ColE1 origin"  
3181..4412  
promoter /note="rice actin 1' region (actl promoter)"  
/citation=[2]

## ORIGIN

Query Match 47.4%; Score 2760.8; DB 12; Length 4412;  
Blast Local Similarity 96.5%; Pred. No. 0;  
Matches 2892; Conservative 0; Mismatches 32; Indels 72; Gaps 4;  
QY 2831 AGTCAAGCAGATCGTTCAAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCTCTGTGCG 2890  
DB 21 ATTCTGCAGATGTTCAACATTTGGCAATAAAGTTTCTTAAGATTGAATCTCTGTGCG 80  
QY 2891 GGTCTTGCAGATGATTATCATATAATTTCTGTGGAATTAGCTTAAAGCATGTAATAATAAC 2950  
DB 81 GGTCTTGCAGATGATTATCATATAATTTCTGTGGAATTAGCTTAAAGCATGTAATAATAAC 140  
QY 2951 ATGTAATGATGACGTTATTATGAGATGCGTTTATGATGATGAGTCCGCGCAATTATAC 3010  
DB 141 ATGTAATGATGACGTTATTATGAGATGCGTTTATGATGATGAGTCCGCGCAATTATAC 200  
QY 3011 ATTTAATACGATAGAAAACAAAATATAGCGCGCAAACTAGSATAAATATCGCGCGG 3070  
DB 201 ATTTAATACGATAGAAAACAAAATATAGCGCGCAAACTAGSATAAATATCGCGCGG 260  
QY 3071 GTGTATCTATGTTACTAGATGACCTGCGAGCATGGGATTCGCGCGCGCGATCGGACGTC 3130  
DB 261 GTGTATCTATGTTACTAGATGACCTGCGAGCGCGCGCGCGCGGAGC----- 310  
QY 3131 GGGCCCAATTCGCCCTATAGTACGTTATTAC---AATTCACTGCGCGCTGCTTTTACAA 3187  
DB 311 ---TCCAAATTCGCCCTATAGTACGTTATTACGCGCGCTCACTGCGCGCTGCTTTTACAA 367  
QY 3188 CGTCGTGACTGGGAAAACCTCGCGCTTACCCAACTTAATCGCTTTCGACACATCCCTCT 3247  
DB 368 CGTCGTGACTGGGAAAACCTCGCGCTTACCCAACTTAATCGCTTTCGACACATCCCTCT 427  
QY 3248 TTCGCCAGCTGGGTAATAGGAGAGGCCCGGACCGATCGCTTCCCAACAGTTGGCG 3307  
DB 428 TTCGCCAGCTGGGTAATAGGAGAGGCCCGGACCGATCGCTTCCCAACAGTTGGCG 487  
QY 3308 AGCCTGAATGGCGAAT-GGACCGCGCTTGTAGCGGCATTAAGCGCGCGGCTGTGGT 3366  
DB 486 AGCCTGAATGGCGAATGGACCGCGCTTGTAGCGGCATTAAGCGCGCGGCTGTGGT 1572

4447 GATCGTTGGGAACCGGAGCTGAATGAAGCATATCCAAAACGACGAGCGTGACACCAAGATG 4506  
1573 GATCGTTGGGAACCGGAGCTGAATGAAGCATATCCAAAACGACGAGCGTGACACCAAGATG 1632  
4507 CCGTAGCAATGCAACAGCTTGGCCCAACTATTAACTGGCCAACTACTTACTCTAGCT 4566  
1633 CCGTAGCAATGCAACAGCTTGGCCCAACTATTAACTGGCCAACTACTTACTCTAGCT 1692  
4567 TCCCGCAACAAATTAATAGACTGGAGCGGATAAAGTTGACGAGCACCTTCTGCGC 4626  
1693 TCCCGCAACAAATTAATAGACTGGAGCGGATAAAGTTGACGAGCACCTTCTGCGC 1752  
4627 TCGGCCCTTCGCGCTCGCTGCTTTTATCTGTGAATAATCTGGAGCCGGTAGCGGTCT 4686  
1753 TCGGCCCTTCGCGCTCGCTGCTTTTATCTGTGAATAATCTGGAGCCGGTAGCGGTCT 1812  
4687 CGCGGTATCAATTCAGCACTGGGCGCAGAGTGAAGCCCTCCGATATCGTGTATCTAC 4746  
1813 CGCGGTATCAATTCAGCACTGGGCGCAGAGTGAAGCCCTCCGATATCGTGTATCTAC 1872  
4747 ACGACGGGAGTCAGGCAACTATGATGAACGAAATAGACAGATCGCTGAGATAGGTGCC 4806  
1873 ACGACGGGAGTCAGGCAACTATGATGAACGAAATAGACAGATCGCTGAGATAGGTGCC 1932  
4807 TCACTGATTAAGCAATGCTGTGACGCAAGTTTACTCATATATATCTTTAGATTTGAT 4866  
1933 TCACTGATTAAGCAATGCTGTGACGCAAGTTTACTCATATATATCTTTAGATTTGAT 1992  
4867 TTAATACTTCAATTTTAAATTTAAAGATCTAGTGAAGATCTTTTGTGATATCTCATG 4926  
1993 TTAATACTTCAATTTTAAATTTAAAGATCTAGTGAAGATCTTTTGTGATATCTCATG 2052  
4927 ACCAAATCCCTTAAAGCTGAGTTTCTGCTCCACTGAGCGTGCAGACCCGCTAGAAAAGATC 4986  
2053 ACCAAATCCCTTAAAGCTGAGTTTCTGCTCCACTGAGCGTGCAGACCCGCTAGAAAAGATC 2112  
4987 AAGGATCTTCTTGAGATCTTTTCTGCGGTAAATCTGCTGCTTGCAGAAACAAAAA 5046  
2113 AAGGATCTTCTTGAGATCTTTTCTGCGGTAAATCTGCTGCTTGCAGAAACAAAAA 2172  
5047 CCACCGCTACCGGCTGTTCTTTCGCGGATCAAGAGCTACCACTCTTTTTCGGAAG 5106  
2173 CCACCGCTACCGGCTGTTCTTTCGCGGATCAAGAGCTACCACTCTTTTTCGGAAG 2232  
5107 GTAACCTGGCTTACGACAGCGCAGATACCAAACTACTGCTCTTCTAGTGTAGCGGTAGTTA 5166  
2233 GTAACCTGGCTTACGACAGCGCAGATACCAAACTACTGCTCTTCTAGTGTAGCGGTAGTTA 2292  
5167 GGCACCACTTCAAGAACTCTGTAGCAACCGCTTACATACCTGCTCTGCTAAATCTGTTA 5226  
2293 GGCACCACTTCAAGAACTCTGTAGCAACCGCTTACATACCTGCTCTGCTAAATCTGTTA 2352  
5227 CAGTGGCTGCTGCCAGTGGCGATAAGTCTGCTTACCGGGTTGGACTCAAGACGATAG 5286  
2353 CAGTGGCTGCTGCCAGTGGCGATAAGTCTGCTTACCGGGTTGGACTCAAGACGATAG 2412  
5287 TTACCGGATAAGCGCGCTGCGGCTGAAACGGGGGTTTGTGCAACACAGCCAGCTTG 5346  
2413 TTACCGGATAAGCGCGCTGCGGCTGAAACGGGGGTTTGTGCAACACAGCCAGCTTG 2472  
5347 GAGCGAACCACTTACACCGAACTGAGATACCTACAGGTGAGCTATGAGAAACGCGCACG 5406  
2473 GAGCGAACCACTTACACCGAACTGAGATACCTACAGGTGAGCTATGAGAAACGCGCACG 2532  
5407 CTTCCGAAAGGGAAGCGGACAGTATCCGTTAAGCGGCGAGGTTCGGAACAGAGAG 5466  
2533 CTTCCGAAAGGGAAGCGGACAGTATCCGTTAAGCGGCGAGGTTCGGAACAGAGAG 2592  
5467 CGCACGAGGAGCTTCCAGGGGAAACGCTGTTATCTTTTATAGTCTCTGCGGGTTTTCG 5526  
2593 CGCACGAGGAGCTTCCAGGGGAAACGCTGTTATCTTTTATAGTCTCTGCGGGTTTTCG 2652  
5527 CACCTCTGACTTGAGCGTGAATTTTGTGATGCTCTGTCAGGGGGGCGAGCCTATCGAAA 5586

2653 CACCTCTGACTTGAGCGTGAATTTTGTGATGCTCGTTCAGGGGGCGAGCCTATGGAAA 2712  
5587 AACGCCAGAAACCGCGCTTTTACCGTTCTTGGCCCTTTTGTGCGCTTTTGTCTCACATG 5646  
2713 AACGCCAGAAACCGCGCTTTTACCGTTCTTGGCCCTTTTGTGCGCTTTTGTCTCACATG 2772  
5647 TTTCTTCTGCTGCTTATCCCTGATTCGTGTGATAAACCGTATTACCGCTTTGAGTGAAGT 5706  
2773 TTTCTTCTGCTGCTTATCCCTGATTCGTGTGATAAACCGTATTACCGCTTTGAGTGAAGT 2832  
5707 GATACCGCTGCGCGAGCGAAGCAACGAGCGGCGAGCGAGTCACTGAGCGAGGAGCGGAA 5766  
2833 GATACCGCTGCGCGAGCGAAGCAACGAGCGGCGAGCGAGTCACTGAGCGAGGAGCGGAA 2892  
5767 GAGCGCCCAATACGCAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAAATGAG 5822  
2893 GAGCGCCCAATACGCAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAAATGAG 2948

RESULT 3  
AR037157  
LOCUS AR037157 5534 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 3 from patent US 5801027.  
ACCESSION AR037157  
VERSION AR037157.1 GI:5955013  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 5534)  
AUTHORS Bennett, M., May, S. and Ramsay, N.  
TITLE Method of using transactivation proteins to control gene expression in transgenic plants  
JOURNAL Patent: US 5801027-A 3 01-SEP-1998;  
FEATURES  
source 1. 5534  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 47.3%; Score 2755.6; DB 6; Length 5534;  
Best Local Similarity 96.3%; Pred. No. 0;  
Matches 2890; Conservative 0; Mismatches 38; Indels 74; Gaps 4;

QY 2840 GATCGTTCAAACTTGGCAATAAAGTTTCTTAAGATTGAATCTGTTGCGGCTTTGCG 2899  
DB 2366 GATCGTTCAAACTTGGCAATAAAGTTTCTTAAGATTGAATCTGTTGCGGCTTTGCG 2425  
QY 2900 ATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAAATTAACATGTATGC 2959  
DB 2426 ATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAAATTAACATGTATGC 2485  
QY 2960 ATGACGTTATTTATGATGAGTGGGTTTTTATGATTAGAGTCCGCAATTTATACATTTAATAC 3019  
DB 2486 ATGACGTTATTTATGATGAGTGGGTTTTTATGATTAGAGTCCGCAATTTATACATTTAATAC 2545  
QY 3020 GCATAGAAAAACAAATATAGCGCGCAAACTAGCAATAAATTTATCGCGCGGTGCTATCT 3079  
DB 2546 GCATAGAAAAACAAATATAGCGCGCAAACTAGCAATAAATTTATCGCGCGGTGCTATCT 2605  
QY 3080 ATGTTACTAGATCGACC-----TGACGGCATGGGATCGCGCGCGCATGC 3124  
DB 2606 ATGTTACTAGATCWNNGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGGGG 2665  
QY 3125 GAGTCGGGCGCAATTCGCCCTATAGTGAAGTCTGATTAC--AATTCACCTGGCGGCTGCTT 3181  
DB 2666 GGCGCGGTACCCCAATTCGCCCTATAGTGAAGTCTGATTACCGCGGCTCACTGGCGGCTGCTT 2725  
QY 3182 TTACAAACGTCGTGACTGGGAAAAACCTGGCGTTACCCAACTTAATCCCTTTCGACGACAT 3241  
DB 2726 TTACAAACGTCGTGACTGGGAAAAACCTGGCGTTACCCAACTTAATCCCTTTCGACGACAT 2785

3242 CCCCCTTTCCGCGAGCTGGCGTAATAGCGAAGAGCCCGGACCGATCGCCCTTCTCCACAG 3301  
Db CCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGCCCGGACCGATCGCCCTTCTCCACAG 2845  
3302 TTGCGAGCCTGAAATGCCGAAT--GGACGCGCCCTGTAGCGCGCATTAAGCGCGCGGGT 3360  
Db TTGCGAGCCTGAAATGCCGAATGGGAGCGCGCCCTGTAGCGCGCATTAAGCGCGCGGGT 2905  
3361 GTGGTGGTTAGCGCGAGCTGAGCGGTACACTTGTCGACGCGCCCTAGCGCGCTCTCTTTTC 3420  
Db GTGGTGGTTAGCGCGAGCTGAGCGGTACACTTGTCGACGCGCCCTAGCGCGCTCTCTTTTC 2965  
3421 GCTTTCTTCCCTTCTCTTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAATCG 3480  
Db GCTTTCTTCCCTTCTCTTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAATCG 3025  
3481 GGCTCCCTTTAGGGTTCGAAATTAGAGCTTTAGCGCACCTCGACCGCAAAACTTGAT 3540  
Db GGCTCCCTTTAGGGTTCGAAATTAGAGCTTTAGCGCACCTCGACCGCAAAACTTGAT 3085  
3026 GGCTCCCTTTAGGGTTCGAAATTAGAGCTTTAGCGCACCTCGACCGCAAAACTTGAT 3085  
3541 TTGGGTGARGGTTCACTAGTGGGCCATCGCCCTGTAGACGGTTTTCGCCCTTTGAOG 3600  
Db TAGGTGATGCTTCACTAGTGGGCCATCGCCCTGTAGACGGTTTTCGCCCTTTGAOG 3145  
3601 TTGGAGTCCAGGTTCTTAAATAGTGGACTCTGTGTTCCAACTGGAAACACACTCAACCT 3660  
Db TTGGAGTCCAGGTTCTTAAATAGTGGACTCTGTGTTCCAACTGGAAACACACTCAACCT 3205  
3146 TTGGAGTCCAGGTTCTTAAATAGTGGACTCTGTGTTCCAACTGGAAACACACTCAACCT 3205  
3661 ATCTCGGTCTATTCTTTGATTTATAGGGAATTTGCGGATTTTCGGCTATTGTTTAAAA 3720  
Db ATCTCGGTCTATTCTTTGATTTATAGGGAATTTGCGGATTTTCGGCTATTGTTTAAAA 3265  
3206 ATCTCGGTCTATTCTTTGATTTATAGGGAATTTGCGGATTTTCGGCTATTGTTTAAAA 3265  
3721 AATGAGCTGATTTAAACAAATATTAAACGCGAATTTTAAACAAATATTAAAGCTTTACAAT 3780  
Db AATGAGCTGATTTAAACAAATATTAAACGCGAATTTTAAACAAATATTAAAGCTTTACAAT 3325  
3781 TCCGCTGATGCGGTATTTCTCTTACGCACTGTGCGGTATTTTCACACGCGATACAGT 3840  
Db T-----AGT 3330  
3841 GGCACTTTTCGGGAAATGCGCGAAACCCCTATTGTTTATTTTCTAAATACATTCA 3900  
Db GGCACTTTTCGGGAAATGCGCGAAACCCCTATTGTTTATTTTCTAAATACATTCA 3390  
3901 AATATGATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATTGAAAGG 3960  
Db AATATGATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATTGAAAGG 3450  
3961 AAGAGTATGAGTATTCAACTTTCCGTGCGCTTATCCCTTTTTCGGGCAATTTGC 4020  
Db AAGAGTATGAGTATTCAACTTTCCGTGCGCTTATTCCTTTTTCGGGCAATTTGC 3510  
4021 CTTCTGTTTTCGCTCACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTG 4080  
Db CTTCTGTTTTCGCTCACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTG 3570  
4081 GGTGACAGTGGGTTTACATGAACTGGAATCTCAACAGCGGTAAAGATCCTTGAGAGTTT 4140  
Db GGTGACAGTGGGTTTACATGAACTGGAATCTCAACAGCGGTAAAGATCCTTGAGAGTTT 3630  
4141 CGCCCGAAGAACGTTTTCATGATGAGACACTTTTAAAGTCTGCTATGTGCGCGGTA 4200  
Db CGCCCGAAGAACGTTTTCATGATGAGACACTTTTAAAGTCTGCTATGTGCGCGGTA 3690  
4201 TTATCCGCTATTACGCGCGGCAAGAGCACTCGGTGCGCGCATACACTATTCTCAGAAT 4260  
Db TTATCCGCTATTACGCGCGGCAAGAGCACTCGGTGCGCGCATACACTATTCTCAGAAT 3750  
4261 GACTTGGTTGAGTACTCACAGTTCACAGAAAGCATCTTACGGATGGCATGACAGTAAGA 4320  
Db GACTTGGTTGAGTACTCACAGTTCACAGAAAGCATCTTACGGATGGCATGACAGTAAGA 3810  
4321 GATATATGAGTGTGCTGCCATACCAATGAGTGATAACACTGCGGCCAATCTTCTTGACA 4380

3811 GAATATGAGTGTGCTGCCATACCAATGAGTGATAACACTGCGGCCAATCTTACTTCTGACA 3870  
Qy ACCATCGGAGGACCGGAGAGCTAAACCGCTTTTTCACAACTATGGGGGATCATGTAACT 4440  
Db ACCATCGGAGGACCGGAGAGCTAAACCGCTTTTTCACAACTATGGGGGATCATGTAACT 3930  
4441 CGCTTTGATCTGTTGGGAAACCGGAGCTGAATGAAGCCATACAAACGACGAGCGTGACACC 4500  
Db CGCTTTGATCTGTTGGGAAACCGGAGCTGAATGAAGCCATACAAACGACGAGCGTGACACC 3990  
4501 ACCATGCTCTAGCAATGSCAAACAACTGTCGCGCAAACTATTAATCTGCGAACTTACT 4560  
Db ACCATGCTCTAGCAATGSCAAACAACTGTCGCGCAAACTATTAATCTGCGAACTTACT 4050  
4561 CTAGCTTTCCGCGCAACAAATTAATAGACTGATGAGCGCGGATAAAGTTGACGACCACTT 4620  
Db CTAGCTTTCCGCGCAACAAATTAATAGACTGATGAGCGCGGATAAAGTTGACGACCACTT 4110  
4621 CTGCGCTCGGCCCTTCCGCTGCTGTTTATTTGCTGATAAATCTGAGCGCGTGAGCGT 4680  
Db CTGCGCTCGGCCCTTCCGCTGCTGTTTATTTGCTGATAAATCTGAGCGCGTGAGCGT 4170  
4681 GGTCTCGCGGTATCATTTGCGCACTGCGGCGCAGATGTTAAGCCCTCCCGTATCGTAGTT 4740  
Db GGTCTCGCGGTATCATTTGCGCACTGCGGCGCAGATGTTAAGCCCTCCCGTATCGTAGTT 4230  
4741 ATCTACACGACGCGGAGTCAAGCACTATGCGATGAACGAAATAGACAGATCGCTGAGATA 4800  
Db ATCTACACGACGCGGAGTCAAGCACTATGCGATGAACGAAATAGACAGATCGCTGAGATA 4290  
4801 GGTGCTCTCATGATTAAGCAATGCTGTAACCTGTCAGACCAAGTTTACTCATATATCTTTAG 4860  
Db GGTGCTCTCATGATTAAGCAATGCTGTAACCTGTCAGACCAAGTTTACTCATATATCTTTAG 4350  
4861 ATTGATTTAAACCTTCAATTTTAAAGATCTTAGTGAGATGCTTTTGTGATAAT 4920  
Db ATTGATTTAAACCTTCAATTTTAAAGATCTTAGTGAGATGCTTTTGTGATAAT 4410  
4921 CTATGACCAAAATTCCTTAAACGAGTATTTTCGTCACCTGAGCGTCAGACCCCGTAGAA 4980  
Db CTATGACCAAAATTCCTTAAACGAGTATTTTCGTCACCTGAGCGTCAGACCCCGTAGAA 4470  
4981 AGATCAAAAGGATCTTCTTGAGATCTTTTTCGCGGCTAATCTGCTGCTGTCGCAACA 5040  
Db AGATCAAAAGGATCTTCTTGAGATCTTTTTCGCGGCTAATCTGCTGCTGTCGCAACA 4530  
5041 AAAAAACCAACCGCTACAGCGGTGTTTTCGCGGATCAAGAGCTACCAACTCTTTT 5100  
Db AAAAAACCAACCGCTACAGCGGTGTTTTCGCGGATCAAGAGCTACCAACTCTTTT 4590  
5101 CCGAAGGTAACTGCGCTTACGAGCGCAGATACCAAACTACTCTCTTCTAGTGTAGCG 5160  
Db CCGAAGGTAACTGCGCTTACGAGCGCAGATACCAAACTACTCTCTTCTAGTGTAGCG 4650  
5161 TAGTTAGGCAACCACTTCAAGAACTCTGAGACCCGCTACATACCTGCTGCTGCTAATC 5220  
Db TAGTTAGGCAACCACTTCAAGAACTCTGAGACCCGCTACATACCTGCTGCTGCTAATC 4710  
5221 CTGTTACCAAGTGTCTGCGGATAGTCTGCTTACCGGTTGAGTCAAGA 5280  
Db CTGTTACCAAGTGTCTGCGGATAGTCTGCTTACCGGTTGAGTCAAGA 4770  
5281 CGATAGTTAACCGGATTAAGCGCGCTGCGGTGAAACCGGGGTTTCTGTCACACAGCCC 5340  
Db CGATAGTTAACCGGATTAAGCGCGCTGCGGTGAAACCGGGGTTTCTGTCACACAGCCC 4830  
5341 AGCTTGGAGGCAACGACTTACCGGAACTGAGTATCTACAGCGTGAAGTATGAGAGC 5400  
Db AGCTTGGAGGCAACGACTTACCGGAACTGAGTATCTACAGCGTGAAGTATGAGAGC 4890  
5401 GGCACGCTTCCGGAAGGAGAAAGCGGACAGGTATCTCGGTAGCGCGAGGTCGGAACA 5460

Db 4891 GCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGACGGGTGCGGAACA 4950  
Qy 5461 GGAAGAGCGCACGAGGGAGCTTCCAGGGGGGAAAGCGCTGGTATCTTTATAGTCTCTGTCCGG 5520  
Db 4951 GGAAGAGCGCACGAGGGAGCTTCCAGGGGGGAAAGCGCTGGTATCTTTATAGTCTCTGTCCGG 5010  
Qy 5521 TTTTCGCACTCTGACCTTGAGAGCTGAGTTTGTGATGCTGTCAGGGGGGCGAGCCTTA 5580  
Db 5011 TTTTCGCACTCTGACCTTGAGAGCTGAGTTTGTGATGCTGTCAGGGGGGCGAGCCTTA 5070  
Qy 5581 TGGAAAAACCCAGACGCGGCTTTTACGGTTCTCTGCGCTTCTGCGCTTTTGTGCTGGCTTTGCT 5640  
Db 5071 TGGAAAAACCCAGACGCGGCTTTTACGGTTCTCTGCGCTTCTGCGCTTTTGTGCTGGCTTTGCT 5130  
Qy 5641 CACATGTTCTTCTGCTTATCCCTGATTTCTGTGATAACCGTATTACCGCTTTTGAG 5700  
Db 5131 CACATGTTCTTCTGCTTATCCCTGATTTCTGTGATAACCGTATTACCGCTTTTGAG 5190  
Qy 5701 TGAGCTGATACCGCTGCGCGAGCGGACGACGAGCGGAGCGAGTCAGTGAAGCGGAGAA 5760  
Db 5191 TGAGCTGATACCGCTGCGCGAGCGGACGACGAGCGGAGCGAGTCAGTGAAGCGGAGAA 5250  
Qy 5761 GCGGAAGAGCGCCCAATAGCAAAACGCTCTCCCGCGGTTGGCGGATTTAATGTC 5820  
Db 5251 GCGGAAGAGCGCCCAATAGCAAAACGCTCTCCCGCGGTTGGCGGATTTAATGTC 5310  
Qy 5821 AG 5822  
Db 5311 AG 5312

RESULT 4  
AR112043  
LOCUS Sequence 3 from patent US 6127606.  
DEFINITION AR112043  
ACCESSION AR112043  
VERSION AR112043.1 GI:12828891  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5534)  
AUTHORS Bennett,M., May,S. and Ramsay,N.  
TITLE Method of using transactivation proteins to control expression in transgenic plants  
JOURNAL Patent: US 6127606-A 3 03-OCT-2000;  
FEATURES  
source Location/Qualifiers  
1..5534  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 47.3%; Score 2755.6; DB 6; Length 5534;  
Best Local Similarity 96.3%; Pred.No. 0;  
Matches 2890; Conservative 0; Mismatches 38; Indels 74; Gaps 4;  
Qy 2840 GATCGTTCAACACTTTGGCAATAAGTTTCTTAAGATTGAATCCTGTGCGGTCTTGCG 2899  
Db 2366 GATCGTTCAACACTTTGGCAATAAGTTTCTTAAGATTGAATCCTGTGCGGTCTTGCG 2425  
Qy 2900 ATGATTATCATATAATTTCTGTGATTGATGTTAAGCTTAAGCTGTAATTAACATGTAATGC 2959  
Db 2426 ATGATTATCATATAATTTCTGTGATTGATGTTAAGCTTAAGCTGTAATTAACATGTAATGC 2485  
Qy 2960 ATGACGTTATTTATGATGAGTGGTTTATGATTAGAGTCCCGCAATTAACATTTAATAC 3019  
Db 2486 ATGACGTTATTTATGATGAGTGGTTTATGATTAGAGTCCCGCAATTAACATTTAATAC 2545  
Qy 3020 GCGATAGAAAACAAATATAGCGCGCAAACTAGGATAAAATATCGCGCGGTGTCTATCT 3079  
Db 2546 GCGATAGAAAACAAATATAGCGCGCAAACTAGGATAAAATATCGCGCGGTGTCTATCT 2605  
Qy 3080 ATGTTACTAGATCGACC-----TGCAGGCATGGGATCCGCGCGCGCATGC 3124

Db 2606 ATGTTACTAGATCANNNGAATTCGATATCAAGCTTATCGATACCGTTCGACCTCGAGGGGG 2665  
Qy 3125 GAGTGGGGCCCAATTCGCCCTATAGTAGTGGTATTAAC---AATTCACTGGCGGTGGTT 3181  
Db 2666 GGGCGGTACCCCAATTCGCCCTATAGTAGTGGTATTAACGGCGCTCACTGGCGGTGGTT 2725  
Qy 3182 TTACAAGCTGCTGACTGGGAAACCCCTGGGGTTACCCCACTTAATCGCTTTGCGACACAT 3241  
Db 2726 TTACAAGCTGCTGACTGGGAAACCCCTGGGGTTACCCCACTTAATCGCTTTGCGACACAT 2785  
Qy 3242 CCGCTTTTCGCGAGCTGGCGTAATAGCGAAAGAGGCGCGACCGATCGCTTCCCAACAG 3301  
Db 2786 CCGCTTTTCGCGAGCTGGCGTAATAGCGAAAGAGGCGCGACCGATCGCTTCCCAACAG 2845  
Qy 3302 TTGCGCAGCTGAATGGCGAAT--GGACGCGCTGTAGCGCGCATTAAGCGCGCGGGT 3360  
Db 2846 TTGCGCAGCTGAATGGCGAATGGGACGCGCTGTAGCGCGCATTAAGCGCGCGGGT 2905  
Qy 3361 GTGGTGGTTACGCGCAGCGTGAACGCTACACTTGCACGCGCCCTAGCGCGCTCTCTTC 3420  
Db 2906 GTGGTGGTTACGCGCAGCGTGAACGCTACACTTGCACGCGCCCTAGCGCGCTCTCTTC 2965  
Qy 3421 GCTTTCTTCCCTTCTTCTTCTGCGCACGTTGCGCGGCTTTCCCGTCAAGCTCTAAATCGG 3480  
Db 2966 GCTTTCTTCCCTTCTTCTTCTGCGCACGTTGCGCGGCTTTCCCGTCAAGCTCTAAATCGG 3025  
Qy 3481 GGGCTCCCTTTAGGGTTCCGATTTAGAGCTTTTACGGCACCTCGACCGCAAAAAAATTTGAT 3540  
Db 3026 GGGCTCCCTTTAGGGTTCCGATTTAGAGCTTTTACGGCACCTCGACCGCAAAAAAATTTGAT 3085  
Qy 3541 TTGGGTGATGGTTCAAGTAGTGGGCCATCCCTCTGATAGACGGTTTTCGCCCTTTGAGC 3600  
Db 3086 TAGGGTGAATGTTTCAAGTAGTGGGCCATCCCTCTGATAGACGGTTTTCGCCCTTTGAGC 3145  
Qy 3601 TTGGAGTCAAGCTTCTTTAATAGTGAAGCTTGTTCGAACTGGGACAACTCAACGCT 3660  
Db 3146 TTGGAGTCAAGCTTCTTTAATAGTGAAGCTTGTTCGAACTGGGACAACTCAACGCT 3205  
Qy 3661 ATCTCGGTCTATTCTTTTGTGATTTAAGGATTTTTCGGATTTTCGGCTATTGTGTTAAAA 3720  
Db 3206 ATCTCGGTCTATTCTTTTGTGATTTAAGGATTTTTCGGATTTTCGGCTATTGTGTTAAAA 3265  
Qy 3721 AATGAGCTGATTTAAACAAATATTTAAACGGAATTTTAAACAAATATTTAAAGTTTACAA 3780  
Db 3266 AATGAGCTGATTTAAACAAATATTTAAACGGAATTTTAAACAAATATTTAAAGTTTACAA 3325  
Qy 3781 TCGCCCTGATCGGGTATTTTCTCCTTACGCACTCTGTGGGTATTTTCACACCGCATACAGT 3840  
Db 3326 T-----AGGT 3330  
Qy 3841 GGCACCTTTTCGGGAAATGTGCGGAAACCCCTATTGTTTATTTTCTAAATACATTCA 3900  
Db 3331 GGCACCTTTTCGGGAAATGTGCGGAAACCCCTATTGTTTATTTTCTAAATACATTCA 3390  
Qy 3901 AATATGATCTCGCTCATGAGACAAATAAACCCTGATAAATGCTCAATAATTTGAAAAAGG 3960  
Db 3391 AATATGATCTCGCTCATGAGACAAATAAACCCTGATAAATGCTCAATAATTTGAAAAAGG 3450  
Qy 3961 AAGAGTATCAGTATTTCAACATTTCCGTGCGCCCTTATTCCTTTTTCGGGCAATTTTCG 4020  
Db 3451 AAGAGTATCAGTATTTCAACATTTCCGTGCGCCCTTATTCCTTTTTCGGGCAATTTTCG 3510  
Qy 4021 CTTCTGTTTTCCTCACAACCGGAAACCGCTGGTGAAGTAAAGATGCTGAAGATCAGTTG 4080  
Db 3511 CTTCTGTTTTCCTCACAACCGGAAACCGCTGGTGAAGTAAAGATGCTGAAGATCAGTTG 3570  
Qy 4081 GGTGACAGAGTGGTTTACATCGAACTGGAATCTCAACAGCGGTGAAGATCTTTGAGAGTTT 4140  
Db 3571 GGTGACAGAGTGGTTTACATCGAACTGGAATCTCAACAGCGGTGAAGATCTTTGAGAGTTT 3630  
Qy 4141 CCGCCCGAAGAGCTTTTCCCAATGATGACACCTTTTAAAGTTCTGCTATCTGTCGCGGCTA 4200



3631 CGCCCCGAGAGAGGTTTTCACATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGGGTA 1690  
4201 TTATCCCGTATTGACCGCCGGGCAAGAGCAACTCGCGTGGCGGATACATATTTCTCAGAAAT 4260  
3691 TTATCCCGTATTGACCGCCGGGCAAGAGCAACTCGGTGGCGGATACATATTTCTCAGAAAT 3750  
4261 GACCTGGTTGAGTACTACACAGTCAAGAGGAGGATCTTACCGATGGCATGACGTAAGA 4320  
3751 GACCTGGTTGAGTACTACACAGTCAAGAGGAGGATCTTACCGATGGCATGACGTAAGA 3810  
4321 GAATTTATGAGTGTGCTGCCATAACCATGAGTGAATACACTCGCGGCCAACTTACTTCTGACA 4380  
3811 GAATTTATGAGTGTGCTGCCATAACCATGAGTGAATACACTCGCGGCCAACTTACTTCTGACA 3870  
4381 AGATCGGAGGACCGAAGGAGTAAACCGCTTTTTCGCAACAATGGGGGATCATGTAAT 4440  
3871 AGATCGGAGGACCGAAGGAGTAAACCGCTTTTTCGCAACAATGGGGGATCATGTAAT 3930  
4441 CGCCTTGATGCTTGGGACCGGAGCTGAATGAAGCCATACCAACGACGAGGCTGACACC 4500  
3931 CGCCTTGATGCTTGGGACCGGAGCTGAATGAAGCCATACCAACGACGAGGCTGACACC 3990  
4501 ACGATGCTGTAGCAATGCGCAACAAAGTTGCGCAAACTATTAACTGGCGAACTACTTACT 4560  
3991 ACGATGCTGTAGCAATGCGCAACAAAGTTGCGCAAACTATTAACTGGCGAACTACTTACT 4050  
4561 CTAGCTTCCCGGCAACAAATTAATAGACTGGAGGCGGATTAAGTTGCGAGAACACTTT 4620  
4051 CTAGCTTCCCGGCAACAAATTAATAGACTGGAGGCGGATTAAGTTGCGAGAACACTTT 4110  
4621 CTGCGCTCGGCTTCCCGGCTGGCTGCTTTTATGCTGATAAATCTGAGCCCGGTGAGCT 4680  
4111 CTGCGCTCGGCTTCCCGGCTGGCTGCTTTTATGCTGATAAATCTGAGCCCGGTGAGCT 4170  
4681 GGGTCTCGCGGTATCATTTGAGCACTGGGCGCAGATGTAAGCCCTCCCGTATCGTAGTT 4740  
4171 GGGTCTCGCGGTATCATTTGAGCACTGGGCGCAGATGTAAGCCCTCCCGTATCGTAGTT 4230  
4741 ATCTACACACCGGAGTACGAGCACTATGAGTGAACGAAATAGACAGATCGCTGAGATA 4800  
4231 ATCTACACACCGGAGTACGAGCACTATGAGTGAACGAAATAGACAGATCGCTGAGATA 4290  
4801 GGTGCTCACTGATTAAGCATTTGTAACCTGTCAGACCAAGTTTACTCATATATACTTTAG 4860  
4291 GGTGCTCACTGATTAAGCATTTGTAACCTGTCAGACCAAGTTTACTCATATATACTTTAG 4350  
4861 ATTGATTTAAAACTTCAATTTTAAATTTAAAGGATCTAGTGAAGATCCTTTTGTATAT 4920  
4351 ATTGATTTAAACTTCATTTTAAATTTAAAGGATCTAGTGAAGATCCTTTTGTATAT 4410  
4921 CTCTAGACCAAAATCCCTTAAAGTGAATTTTGTTCCTCACTGAGCGTCAGACCCGTAGAA 4980  
4411 CTCTAGACCAAAATCCCTTAAAGTGAATTTTGTTCCTCACTGAGCGTCAGACCCGTAGAA 4470  
4981 AAGATCAAGGATCTTCTGAGATCCTTTTCTGCGGTAATCTGCTGCTTGCAACA 5040  
4471 AAGATCAAGGATCTTCTGAGATCCTTTTCTGCGGTAATCTGCTGCTTGCAACA 4530  
5041 AAAAAACCAACCGCTACACGCGGTGTTGTTTCCCGGATCAAGAGCTACCAACTCTTTT 5100  
4531 AAAAAACCAACCGCTACACGCGGTGTTGTTTCCCGGATCAAGAGCTACCAACTCTTTT 4590  
5101 CCGAAGTTAACTGGCTTACGAGGAGCGGATACCAATACTGCTCTTCTAGTGTAGCG 5160  
4591 CCGAAGTTAACTGGCTTACGAGGAGCGGATACCAATACTGCTCTTCTAGTGTAGCG 4650  
5161 TAGTTAGGCAACACTTCAAGAACTCTGTAGCAACCGCTTACATCTGCTCTGCTAATC 5220  
4651 TAGTTAGGCAACACTTCAAGAACTCTGTAGCAACCGCTTACATCTGCTCTGCTAATC 4710  
5221 CTGTTACCAAGTGGCTGCTCCAGTGGCGATAAGTCTGTTCTTACCGGGTTGACTCAAGA 5280  
4711 CTGTTACCAAGTGGCTGCTCCAGTGGCGATAAGTCTGTTCTTACCGGGTTGACTCAAGA 4770

5281 CGATAGTTACCGGATAGCGCGCAGCGGTGCGGCTGAACCGGGGGTTCGTGCAACAGCCC 5340  
4771 CGATAGTTACCGGATAGCGCGCAGCGGTGCGGCTGAACCGGGGGTTCGTGCAACAGCCC 4830  
5341 ACCTTGGAGCGAAACGACCTTACACCGAACTGAGATACCTTACAGCGTCTGAGCTTATGAGAAAGC 5400  
4831 ACCTTGGAGCGAAACGACCTTACACCGAACTGAGATACCTTACAGCGTCTGAGCTTATGAGAAAGC 4890  
5401 GCCACGCTTCCCGAAGGAGGAGGAGCGGACAGGTATCCGGTAAAGCGGAGGCTCGGAACA 5460  
4891 GCCACGCTTCCCGAAGGAGGAGGAGGAGCGGACAGGTATCCGGTAAAGCGGAGGCTCGGAACA 4950  
5461 GGAGAGCGCACAGGAGGAGCTTCAAGGGGAAACCGCTGCTATCTTTATAGTCTCTGCGG 5520  
4951 GGAGAGCGCACAGGAGGAGCTTCAAGGGGAAACCGCTGCTATCTTTATAGTCTCTGCGG 5010  
5521 TTTTCCGCACTCTGACTCTGAGCGGTGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTA 5580  
5011 TTTTCCGCACTCTGACTCTGAGCGGTGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTA 5070  
5581 TGGAAAAACCGCAGCAACCGCGCTTTTTCACGCTTCTGCGCTTTTGTGCGCTTTTGTCT 5640  
5071 TGGAAAAACCGCAGCAACCGCGCTTTTTCACGCTTCTGCGCTTTTGTGCGCTTTTGTCT 5130  
5641 CACATGTTCTTCTGCGCTTATCCCTGATTTCTGTGATTAACCGTATTAACCGCTTTGAG 5700  
5131 CACATGTTCTTCTGCGCTTATCCCTGATTTCTGTGATTAACCGTATTAACCGCTTTGAG 5190  
5701 TGAGCTGATACCGCTCGCGCAGCGCAACGACCGAGCGCAGCGAGTCACTGAGCGAGGAA 5760  
5191 TGAGCTGATACCGCTCGCGCAGCGCAACGACCGAGCGCAGCGAGTCACTGAGCGAGGAA 5250  
5761 GCGAAGAGCGCCCAATACGCAACCGCTCTCCCGCGGTTGGCCGATTCATTAATGC 5820  
5251 GCGAAGAGCGCCCAATACGCAACCGCTCTCCCGCGGTTGGCCGATTCATTAATGC 5310  
5821 AG 5822  
5311 AG 5312

RESULT 5  
AF173954/c  
LOCUS AF173954  
DEFINITION Cloning vector pGEM-URA3, complete sequence.  
ACCESSION AF173954  
VERSION AF173954.1 GI:6002958  
KEYWORDS Cloning vector pGEM-URA3  
SOURCE Cloning vector pGEM-URA3  
ORGANISM Cloning vector pGEM-URA3  
artificial sequences; vectors.  
REFERENCE 1 (bases 1 to 4350)  
AUTHORS Wilson,R.B., Davis,D. and Mitchell,A.P.  
TITLE Rapid hypothesis testing with Candida albicans through gene  
disruption with short homology regions  
JOURNAL J. Bacteriol. 181 (6), 1868-1874 (1999)  
MEDLINE 99173911  
PUBMED 10074081  
REFERENCE 2 (bases 1 to 4350)  
AUTHORS Wilson,R.B., Davis,D. and Mitchell,A.P.  
TITLE Direct Submission  
JOURNAL Submitted (02-AUG-1999) Microbiology, Columbia University, 701 W.  
168th St., New York, NY 10032, USA  
FEATURES  
source  
1..4350  
/organism="Cloning vector pGEM-URA3"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:101832"  
/lab\_host="Candida albicans SC5314"  
/note="Candida albicans URA3 gene cloned into pGEM-T"  
3409..4221  
/gene="URA3"





1065 GCAACAAAAAACCCAGCTACAGCGGTTGGTTTTCGCCGATCAGAGCTACAC 1006  
QY TCTTTTTCGGAAGTAACTGGCTTCAGCAGAGCGAGATACCAAAATCTGTCTCTTAGT 5153  
Db TCTTTTTCGGAAGTAACTGGCTTCAGCAGAGCGAGATACCAAAATCTGTCTCTTAGT 946  
QY GTAGCGGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCTCT 5213  
Db GTAGCGGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCTCT 886  
QY GCTAATCTCTGTACCACTGCTGCTGCGAGTAACTGCTGTCTTACCGGTTGGA 5273  
Db GCTAATCTCTGTACCACTGCTGCTGCGAGTAACTGCTGTCTTACCGGTTGGA 826  
QY CTCGAAGACGATAGTTACCGGATAGGCGCAGCGGTGGGCTGAACCGGGGTTCTGTCAC 5333  
Db CTCGAAGACGATAGTTACCGGATAGGCGCAGCGGTGGGCTGAACCGGGGTTCTGTCAC 766  
QY ACAGCCAGCTTGGAGGAAACGACCTACACCGAACTGAGATACCTACAGCGTGAAGTATG 5393  
Db ACAGCCAGCTTGGAGGAAACGACCTACACCGAACTGAGATACCTACAGCGTGAAGTATG 706  
QY AGAAGCGCACCGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGTAAGCGGCGAGGT 5453  
Db AGAAGCGCACCGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGTAAGCGGCGAGGT 646  
QY CGGAACGAGGAGCGCACGAGGAGCTTCCAGGGGGAACCGCTGTATCTTTATAGTCC 5513  
Db CGGAACGAGGAGCGCACGAGGAGCTTCCAGGGGGAACCGCTGTATCTTTATAGTCC 586  
QY TGTGCGGTTTGGCAGCTCTGACTGAGCGTGAATTTTGTGATGCTGTCAGCGGGGCG 5573  
Db TGTGCGGTTTGGCAGCTCTGACTGAGCGTGAATTTTGTGATGCTGTCAGCGGGGCG 526  
QY GAGCCTATGAAAAACCCAGCAACCGCGCTTTTACGGTCTCTGCGCTTTTGTGCGCC 5633  
Db GAGCCTATGAAAAACCCAGCAACCGCGCTTTTACGGTCTCTGCGCTTTTGTGCGCC 466  
QY TTTTGTGCATGATTTCTTCTCGGTTATCCCTGATTTCTGTGATAACCGTATTACCGC 5693  
Db TTTTGTGCATGATTTCTTCTCGGTTATCCCTGATTTCTGTGATAACCGTATTACCGC 406  
QY CTTTGTAGTGAAGTATACCGCTGCGCAGCGCAACCGAGCGAGCGAGTCAAGTGAAG 5753  
Db CTTTGTAGTGAAGTATACCGCTGCGCAGCGCAACCGAGCGAGCGAGTCAAGTGAAG 346  
QY CGAGGAGCGGAAGAGCGCCCAATACCAACCGCTCTCCCGCGGTTGGCGGATTCA 5813  
Db CGAGGAGCGGAAGAGCGCCCAATACCAACCGCTCTCCCGCGGTTGGCGGATTCA 286  
QY TTAATGCAG 5822  
Db TTAATGCAG 277

RESULT 6  
AF173955/c  
LOCUS AF173955 5558 bp DNA circular SYN 01-OCT-1999  
DEFINITION Cloning vector pGEM-HIS1, complete sequence.  
ACCESSION AF173955  
VERSION AF173955.1 GI:6002960  
KEYWORDS Cloning vector pGEM-HIS1  
SOURCE Cloning vector pGEM-HIS1  
ORGANISM artificial sequences; vectors.  
REFERENCE 1 (bases 1 to 5558)  
AUTHORS Wilson, R.B., Davis, D. and Mitchell, A.P.  
TITLE Rapid hypothesis testing with Candida albicans through gene  
disruption with short homology regions  
J. Bacteriol. 181 (6), 1868-1874 (1999)  
MEDLINE 99173911  
PUBMED 14074081

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

2 (bases 1 to 5558)  
Wilson, R.B., Davis, D. and Mitchell, A.P.  
Direct Submission  
Submitted (02-AUG-1999) Microbiology, Columbia University, 701 W.  
168th St., New York, NY 10032, USA  
Location/Qualifiers  
1. 5558  
/organism="Cloning vector pGEM-HIS1"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:101833"  
/lab\_host="Candida albicans SC5314"  
/note="Candida albicans HIS1 gene into pGEM-T"  
2980. 5558  
/gene="HIS1"  
3514. 4410  
/gene="HIS1"  
/codon\_start=1  
/product="ATP phosphoribosyltransferase"  
/protein\_id="AF00227.1"  
/db\_xref="GI:6002961"  
/translation="MDLVNHLPRLLPAPFKGRLYEKCNLLSGADIQFRSRLDI  
ALSTNIPALIFLPAADIPFVPEGNCDLIGLQIKBAEQFDNIEDLDDKPGSK  
LQIQVADVESEKPEQLVGVKIVSFTKLSDFKQSKDPKPMVNIIVORLQGVLA  
VADAVDLVESGETMAAGLKAJETILETSAHLISSKSKPPEMNIIVORLQGVLA  
OEYVLCNYPKSIQAKCLTITPERRAATVSLDKHSDDEEDVAISSMVRKEIGNV  
MDELKAGATDILVLSINCRV"

ORIGIN

Query Match 46.5%; Score 2705.8; DB 12; Length 5558;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 3114 CGCGCGATCGCAGCGTCGGGCCCAATTCGCCCTATAGTGTAGTGTATACAAATTCACCTGG 3173  
Db |||||  
QY 2960 CGCGAGATCGCAGCGTCGGGCCCAATTCGCCCTATAGTGTAGTGTATACAAATTCACCTGG 2901  
Db |||||  
QY 3174 CCGTGTCTTTACACGTCGTGACTGGGAAACCTCTGCGCTTACCCTCACTTAATCGCCTTG 3233  
Db |||||  
QY 2900 CCGTGTCTTTACACGTCGTGACTGGGAAACCTCTGCGCTTACCCTCACTTAATCGCCTTG 2841  
Db |||||  
QY 3234 CAGCACAATCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCTT 3293  
Db |||||  
QY 2840 CAGCACAATCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCTT 2781  
Db |||||  
QY 3294 CCACAAGTTGCGCAGCTGAATGCGGAATGAGAGCGCCCTGTAGCGCGCATTAAGGCG 3353  
Db |||||  
QY 2780 CCACAAGTTGCGCAGCTGAATGCGGAATGAGAGCGCCCTGTAGCGCGCATTAAGGCG 2721  
Db |||||  
QY 3354 GCGGGGTGCTGTGTACGCGCAGCGTACCGCTACACTTGCAGCGCCCTAGCGCCCGC 3413  
Db |||||  
QY 2720 GCGGGGTGCTGTGTACGCGCAGCGTACCGCTACACTTGCAGCGCCCTAGCGCCCGC 2661  
Db |||||  
QY 3414 TCCTTTGCGCTTCTTCCCTTTCTCGCACGTTTCGCCGGCTTTCCCGTCAAGCTCT 3473  
Db |||||  
QY 2660 TCCTTTGCGCTTCTTCCCTTTCTCGCACGTTTCGCCGGCTTTCCCGTCAAGCTCT 2601  
Db |||||  
QY 3474 AAATCGGGGCTCGCTTTAGGTTTCGATTTAGAGCTTTACGGCACTCGACCCGCAAAA 3533  
Db |||||  
QY 2600 AAATCGGGGCTCGCTTTAGGTTTCGATTTAGAGCTTTACGGCACTCGACCCGCAAAA 2541  
Db |||||  
QY 3534 ACTTGATTTGGGTGATGTTTACGAGTGGCCCATCGCCCTGATAGAGCGGTTTTCGCCC 3593  
Db |||||  
QY 2540 ACTTGATTTGGGTGATGTTTACGAGTGGCCCATCGCCCTGATAGAGCGGTTTTCGCCC 2481  
Db |||||  
QY 3594 TTTGACGTTGGAGTCCAAGTCTTTTATAGTGGACTCTTGTTCAAAACCTGGAACAACACT 3653  
Db |||||  
QY 2480 TTTGACGTTGGAGTCCAAGTCTTTTATAGTGGACTCTTGTTCAAAACCTGGAACAACACT 2421  
Db |||||  
QY 3654 CAACCTATCTCGGTCTATTTCTTTTATAGTGGAGTTTTCGCGATTCGCGCTATTG 3713  
Db |||||  
QY 2420 CAACCTATCTCGGTCTATTTCTTTTATAGTGGAGTTTTCGCGATTCGCGCTATTG 2361  
Db |||||  
QY 3714 GTTAAAAAATGAGCTGATTTTAAACAAATATTAAACGCAATTTTAAACAAATATTAAACGCTT 3773  
Db |||||

Db 2360 GTTAAAAATGAGCTGATTTAAACAAATATTTAAACGGAATTTTAAACAAATATTTAAACGTT 2301  
Qy 3774 TACAAATTTCCGCTGATGCGGTATTTCTCTTTACGATCTGCGGTATTTTACACACGCA 3833  
Db 2300 TACAAATTTCCGCTGATGCGGTATTTCTCTTTACGATCTGCGGTATTTTACACACGCA 2241  
Qy 3834 TACAGGTGGCACTTTTCCGGGAAATGTGCGGAAACCCCTATTGTTTAAATTTTCTAAAT 3893  
Db 2240 TACAGGTGGCACTTTTCCGGGAAATGTGCGGAAACCCCTATTGTTTAAATTTTCTAAAT 2181  
Qy 3894 ACATTCAAATATGATTCGGCTCATGAGACAATTAACCCCTGATAAATGCTTCAATATATTG 3953  
Db 2180 ACATTCAAATATGATTCGGCTCATGAGACAATTAACCCCTGATAAATGCTTCAATATATTG 2121  
Qy 3954 AAAAAGGAAGATGAGTATTCACAACTTTCCGCTGTGCGCCCTTATTCCTTTTGGCGG 4013  
Db 2120 AAAAAGGAAGATGAGTATTCACAACTTTCCGCTGTGCGCCCTTATTCCTTTTGGCGG 2061  
Qy 4014 ATTTTGGCTTCTCTGTTTGTCTCACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGA 4073  
Db 2060 ATTTTGGCTTCTCTGTTTGTCTCACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGA 2001  
Qy 4074 TCAGTTGGGTGCAAGAGTGGTTATCATCGAATCTGAATCTCAACAGCGGTAAAGTCTTTGA 4133  
Db 2000 TCAGTTGGGTGCAAGAGTGGTTATCATCGAATCTGAATCTCAACAGCGGTAAAGTCTTTGA 1941  
Qy 4134 GAGTTTTCGCCCCGAGAAAGCTTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATGCG 4193  
Db 1940 GAGTTTTCGCCCCGAGAAAGCTTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATGCG 1881  
Qy 4194 CGCGGTATTTATCCGTTATTCAGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTC 4253  
Db 1880 CGCGGTATTTATCCGTTATTCAGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTC 1821  
Qy 4254 TCAGAAATGATTTGTTGAGTACTGACAGTCAAGAAAGCATTTTACGGATGGCATGAC 4313  
Db 1820 TCAGAAATGATTTGTTGAGTACTGACAGTCAAGAAAGCATTTTACGGATGGCATGAC 1761  
Qy 4314 AGTAAGAGAATTTAGCAGTGTGCTGCAATACCACTGAGTGATTAACCTGCGGCAACTTACT 4373  
Db 1760 AGTAAGAGAATTTAGCAGTGTGCTGCAATACCACTGAGTGATTAACCTGCGGCAACTTACT 1701  
Qy 4374 TCTGACAACTGAGGAGCAAGAGGACTTAACCGCTTTTTCGACAAACATGGGGGATCA 4433  
Db 1700 TCTGACAACTGAGGAGCAAGAGGACTTAACCGCTTTTTCGACAAACATGGGGGATCA 1641  
Qy 4434 TGTAACTCGCCTTGATCTGGGAAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCG 4493  
Db 1640 TGTAACTCGCCTTGATCTGGGAAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCG 1581  
Qy 4494 TGACACCAAGATGCTGTAGCAATGGCAACAACTGTCGCAAACTATTAACTGGCGAACT 4553  
Db 1580 TGACACCAAGATGCTGTAGCAATGGCAACAACTGTCGCAAACTATTAACTGGCGAACT 1521  
Qy 4554 ACTTACTCTAGCTTCCCGGCAACAAATTAATAGACTGAGTGGAGCGGATAAAGTTTCAGG 4613  
Db 1520 ACTTACTCTAGCTTCCCGGCAACAAATTAATAGACTGAGTGGAGCGGATAAAGTTTCAGG 1461  
Qy 4614 ACCACTTCTGGCTCGGCTTCCGCGCTGGCTGTTTATTTGCTGATAAATCTGGAGCCGG 4673  
Db 1460 ACCACTTCTGGCTCGGCTTCCGCGCTGGCTGTTTATTTGCTGATAAATCTGGAGCCGG 1401  
Qy 4674 TGAGCGGTGCTCGGCTGATCATTTGACGACTGCGGCGAGATGTTAAAGCCCTCCCGTAT 4733  
Db 1400 TGAGCGGTGCTCGGCTGATCATTTGACGACTGCGGCGAGATGTTAAAGCCCTCCCGTAT 1341  
Qy 4734 CGTAGTTATCTACACGAGGGGAGTTCAGGCAACTATGATGAACAAATAGACAGATCGC 4793  
Db 1340 CGTAGTTATCTACACGAGGGGAGTTCAGGCAACTATGATGAACAAATAGACAGATCGC 1281  
Qy 4794 TGAGATAGGTGCTCATGATTAAGCAATTTGATGCTGACGCAAGTTTACTCATATAT 4853

Db 1280 TGAGATAGGTGCTCATGATTAAGCAATTTGATTAAGCAATTTTACTCATATAT 1221  
Qy 4854 ACTTTAGATTTGATTTAAAACTTCATTTTAAATTTAAAAAGGATCTAGGTGAAGATCCTTTT 4913  
Db 1220 ACTTTAGATTTGATTTAAAACTTCATTTTAAATTTAAAAAGGATCTAGGTGAAGATCCTTTT 1161  
Qy 4914 TGTAAATCTCATGACCAAAATCCCTTAAACGTGAGTTTTCGTTTCCACTGAGCGTCAGACCC 4973  
Db 1160 TGTAAATCTCATGACCAAAATCCCTTAAACGTGAGTTTTCGTTTCCACTGAGCGTCAGACCC 1101  
Qy 4974 CGTAGAAAAGATCAAGGATCTTTTGGAGATCTTTTCTGCGCGTAATCTGCTGCTT 5033  
Db 1100 CGTAGAAAAGATCAAGGATCTTTTGGAGATCTTTTCTGCGCGTAATCTGCTGCTT 1041  
Qy 5034 GCAAAACAAAACCAACCGCTTACAGCGGTGTTTGTTCGCGGATCAAGACTACCAAC 5093  
Db 1040 GCAAAACAAAACCAACCGCTTACAGCGGTGTTTGTTCGCGGATCAAGACTACCAAC 981  
Qy 5094 TCTTTTTCGGAAGTAACTGGCTTACGACAGCGGAGATACCAATATCTGTCTTCTTAGT 5153  
Db 980 TCTTTTTCGGAAGTAACTGGCTTACGACAGCGGAGATACCAATATCTGTCTTCTTAGT 921  
Qy 5154 GTAGCGGTAGTTAGCCCACTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCTCT 5213  
Db 920 GTAGCGGTAGTTAGCCCACTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCTCT 861  
Qy 5214 GCTAATCTGTTTACAGTGGCTGCTGCGAGTGGCGATTAAGTCTGTCTTTACCGGGTTGA 5273  
Db 860 GCTAATCTGTTTACAGTGGCTGCTGCGAGTGGCGATTAAGTCTGTCTTTACCGGGTTGA 801  
Qy 5274 CTCAGAGCATAGTTACCGGATTAAGCGCGTGGGCTGGAACGGGGGTTCTGTCAC 5333  
Db 800 CTCAGAGCATAGTTACCGGATTAAGCGCGTGGGCTGGAACGGGGGTTCTGTCAC 741  
Qy 5334 ACAGCCAGCTTTCAGAGCAAGCACTACACCGAACTGAGATACCTTACAGCGTGAAGT 5393  
Db 740 ACAGCCAGCTTTCAGAGCAAGCACTACACCGAACTGAGATACCTTACAGCGTGAAGT 681  
Qy 5394 AGAAGCGCCACGCTTCCGGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGGT 5453  
Db 680 AGAAGCGCCACGCTTCCGGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGGT 621  
Qy 5454 CGGAAACAGGAGAGCGACAGGGAGCTTTCAGAGGGGAAAACGGCTGCTTATTTAGTCC 5513  
Db 620 CGGAAACAGGAGAGCGACAGGGAGCTTTCAGAGGGGAAAACGGCTGCTTATTTAGTCC 561  
Qy 5514 TGTGCGGGTTTCGCGCACTCTGACTTTCAGCGTCTGATTTTGTGATGCTCTGTCAGGGGGCG 5573  
Db 560 TGTGCGGGTTTCGCGCACTCTGACTTTCAGCGTCTGATTTTGTGATGCTCTGTCAGGGGGCG 501  
Qy 5574 GAGCCTATGGAAGAAACGCGAGCAACCGCGCTTTTACGGTTCTGCGCTTTTGTGCGC 5633  
Db 500 GAGCCTATGGAAGAAACGCGAGCAACCGCGCTTTTACGGTTCTGCGCTTTTGTGCGC 441  
Qy 5634 TTTTGTCTCATGTTCTTTTCTGCGTATTCCTGATTTCTGTGATTAACCGTATTACCGC 5693  
Db 440 TTTTGTCTCATGTTCTTTTCTGCGTATTCCTGATTTCTGTGATTAACCGTATTACCGC 381  
Qy 5694 CTTTGTAGTGAAGTGAATGCTGCTGCGCGAGCGAAACCGAGCGAGCGAGTCAAGTGAAG 5753  
Db 380 CTTTGTAGTGAAGTGAATGCTGCTGCGCGAGCGAAACCGAGCGAGCGAGTCAAGTGAAG 321  
Qy 5754 CGAGAAACGGAAGAGCGCCCAATACGAAACCGCTCTTCCCGCGGCTGTTGCGCGATTCA 5813  
Db 320 CGAGAAACGGAAGAGCGCCCAATACGAAACCGCTCTTCCCGCGGCTGTTGCGCGATTCA 261  
Qy 5814 TTAATGCAG 5822  
Db 260 TTAATGCAG 252

RESULT 7  
E83777/c

LOCUS B63777 4229 bp DNA linear PAT 27-AUG-2002  
DEFINITION Method for assaying the stability of genetic information in animal  
cells.  
ACCESSION B63777  
VERSION B63777.1 GI:22553615  
KEYWORDS JP 2001087000-A/5.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE 1 (bases 1 to 4229)  
Nakanishi, M. and Ikawa, Y.  
AUTHORS Method for assaying the stability of genetic information in animal  
TITLE Patent: JP 2001087000-A 5 03-APR-2001;  
JOURNAL KK SENTAN KAGAKU GIJUTSU INCUBATION CENTER  
COMMENT OS Artificial Sequence  
FN JP 2001087000-A/5  
PD 03-APR-2001  
PF 17-SEP-1999 JP 1999264320  
PI MASATO NAKANISHI, YUMI IKAWA  
PC C12Q1/68, C12N5/10, C12N15/09, C12Q1/04//A61K48/00, C12N5/00; PC  
C12N15/00  
CC Restriction enzyme BamHI-cut site  
CC Restriction enzyme SmaI-cut site  
FH key Location/Qualifiers.  
FEATURES  
source l. 4229  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
Query Match 46.1%; Score 2686; DB 6; Length 4229;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2686; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3137 AATTGCCCTATAGTCAGTCGTATTACAATTCACTGCGCGTCTGTTTACAAAGTCGTGAC 3196  
DB AATTGCCCTATAGTCAGTCGTATTACAATTCACTGCGCGTCTGTTTACAAAGTCGTGAC 3098  
QY 3197 TGGGAAAACCTTGGCGTTACCCCACTTAATTCGCTTTGCGACACATCCCCCTTTTGGCCAGC 3256  
DB TGGGAAAACCTTGGCGTTACCCCACTTAATTCGCTTTGCGACACATCCCCCTTTTGGCCAGC 3038  
QY 3257 TGGCGTAATAGCGAAGAGCGCCCGACCGATCGCCCTTCCACAGTTGGCGAGCTGAAT 3316  
DB TGGCGTAATAGCGAAGAGCGCCCGACCGATCGCCCTTCCACAGTTGGCGAGCTGAAT 2978  
QY 3317 GGCGAATGGAACGGCCCTGTAGCGGCGCATTAAGCGCGCGGTGAGTGTTACCGCA 3376  
DB GGCGAATGGAACGGCCCTGTAGCGGCGCATTAAGCGCGCGGTGAGTGTTACCGCA 2918  
QY 3377 GGTGACCGGTACACTTGGCAGCGCCCTAGCGCGCGCTCTTTCGCTTTTCCCTTCCT 3436  
DB GGTGACCGGTACACTTGGCAGCGCCCTAGCGCGCGCTCTTTCGCTTTTCCCTTCCT 2858  
QY 3437 TTCTCGCAGCTGCGCGCTTCCCGCTCAAGCTCTTAATCGGCGCTCCCTTAGGGT 3496  
DB TTCTCGCAGCTGCGCGCTTCCCGCTCAAGCTCTTAATCGGCGCTCCCTTTAGGGT 2798  
QY 3497 TCCGATTTAGAGCTTTTACGGCCTCGACCGCAAAAAAATTGATTTGGGTGATGTTTCAC 3556  
DB TCCGATTTAGAGCTTTTACGGCCTCGACCGCAAAAAAATTGATTTGGGTGATGTTTCAC 2738  
QY 3557 GTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTTGACGTTGGAGTCCACGTTCT 3616  
DB GTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTTGACGTTGGAGTCCACGTTCT 2678  
QY 3617 TTAATAGTGGACTTTGTTCCAACTGGGAACAACACTCAACCTATCTCGGTCTATTCTT 3676  
DB TTAATAGTGGACTTTGTTCCAACTGGGAACAACACTCAACCTATCTCGGTCTATTCTT 2618  
QY 3677 TTGATTTAAGGGATTTTTCGGATTTTCGGCTATTGGTTTAAAAAATAGCTGATTTAAC 3736  
DB TTGATTTAAGGGATTTTTCGGATTTTTCGGCTATTGGTTTAAAAAATAGCTGATTTAAC 3736

2617 TTGATTTAAGGGATTTTTCGGATTTTTCGGCTATTGGTTTAAAAAATAGCTGATTTAAC 2558  
QY 3737 AATAATTTAAACGCAATTTTAAACAAATATTAACAGTTTAAACATTTCCCTCGATCGGTAT 3796  
DB AATAATTTAAACGCAATTTTAAACAAATATTAACAGTTTAAACATTTCCCTCGATCGGTAT 2498  
QY 3797 TTCTCTCTTAAGCATCTGTGGGTATTTTACACGGCATACAGTGCGACACTTTTCGGGGA 3856  
DB TTCTCTCTTAAGCATCTGTGGGTATTTTACACGGCATACAGTGCGACACTTTTCGGGGA 2438  
QY 3857 ATGTGCGCGGAACCCCTATTGTTTAAATATTTTAAATATCAATTAATATGATCCGCTCA 3916  
DB ATGTGCGCGGAACCCCTATTGTTTAAATATTTTAAATATCAATTAATATGATCCGCTCA 2378  
QY 3917 TGAGACAATAACCTCGATTAATGCTTCAATATTAATTAAGAAAGAGAGATATGATATTC 3976  
DB TGAGACAATAACCTCGATTAATGCTTCAATATTAATTAAGAAAGAGAGATATGATATTC 2318  
QY 3977 AACATTTCCGTGTGCGCTTATTTCCCTTTTTCGGCATTTTCGGCTTCTGTTTTCCTC 4036  
DB AACATTTCCGTGTGCGCTTATTTCCCTTTTTCGGCATTTTCGGCTTCTGTTTTCCTC 2258  
QY 4037 ACCCAGAAACCGCTGTGAAAGTAAAGATGCTCAAGATCAAGTTGGGTGCGACGATGGTT 4096  
DB ACCCAGAAACCGCTGTGAAAGTAAAGATGCTCAAGATCAAGTTGGGTGCGACGATGGTT 2198  
QY 4097 ACATCGAACTGGATCTCAACAGCGGTAAAGATCTCTGAGAGTCTTTCGGCCCGAAGACGTT 4156  
DB ACATCGAACTGGATCTCAACAGCGGTAAAGATCTCTGAGAGTCTTTCGGCCCGAAGACGTT 2138  
QY 4157 TTCCAATGATGACACTTTTAAAGTTCTGTATGTGGCGGGTATTATCCGCTATTGACG 4216  
DB TTCCAATGATGACACTTTTAAAGTTCTGTATGTGGCGGGTATTATCCGCTATTGACG 2078  
QY 4217 CCGGCAAGAGCAACTCGGTGCGCGCATACACTATTTCTCAGAATGACTTGGTTGAGTACT 4276  
DB CCGGCAAGAGCAACTCGGTGCGCGCATACACTATTTCTCAGAATGACTTGGTTGAGTACT 2018  
QY 4277 CACAGTCAAGAAAAGCATCTTACCGGATGGCATGACAGTAAGAGAAATTAATGCAAGTCTG 4336  
DB CACAGTCAAGAAAAGCATCTTACCGGATGGCATGACAGTAAGAGAAATTAATGCAAGTCTG 1958  
QY 4337 CCATACCATGATGATTAACACTTTCGCGCAACACTTCTTGAACAACATCGAGGAGCCGA 4396  
DB CCATACCATGATGATTAACACTTTCGCGCAACACTTCTTGAACAACATCGAGGAGCCGA 1898  
QY 4397 AGGAGCTAACCGCTTTTTCGCAACAATGGGGGATCATGTAACTCGCTTTCGTTGGG 4456  
DB AGGAGCTAACCGCTTTTTCGCAACAATGGGGGATCATGTAACTCGCTTTCGTTGGG 1838  
QY 4457 AACCGAGCTGAATGAAGCCATACCAACGACGAGGTGACACCAAGTCCCTGTAGCAA 4516  
DB AACCGAGCTGAATGAAGCCATACCAACGACGAGGTGACACCAAGTCCCTGTAGCAA 1778  
QY 4517 TGGCAACAAGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAAC 4576  
DB TGGCAACAAGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAAC 1718  
QY 4577 AATTAATAGCTGAGTGGAGGGATTAAGTTTGCAGGACCACTTCTCGCTTCGCGCTTC 4636  
DB AATTAATAGCTGAGTGGAGGGATTAAGTTTGCAGGACCACTTCTCGCTTCGCGCTTC 1658  
QY 4637 CCGCTGGCTGTTTATTTGCTGATAAATCTGAGCGCGGTGAGCGTGGGTCTCGCGGTATCA 4696  
DB CCGCTGGCTGTTTATTTGCTGATAAATCTGAGCGCGGTGAGCGTGGGTCTCGCGGTATCA 1598  
QY 4697 TTGCAAGCTGGGGCCAGATGTTAGCCCTCCCGTATCGTAGTTATCTACACACCGGGA 4756  
DB TTGCAAGCTGGGGCCAGATGTTAGCCCTCCCGTATCGTAGTTATCTACACACCGGGA 1538  
QY 4757 GTGAGCAACTATGATGAACGAAATAGACAGATCGCTGAGATAGTGGCTCTCATGATTA 4816  
DB GTGAGCAACTATGATGAACGAAATAGACAGATCGCTGAGATAGTGGCTCTCATGATTA 1478













QY 4014 ATTTTGGCTTCTGTTTGTCTACCCAGAAACGCTGCTGTAAGTAAAGATGCTGAAGA 4073  
DB 2085 ATTTTGGCTTCTGTTTGTCTACCCAGAAACGCTGCTGTAAGTAAAGATGCTGAAGA 2026  
QY 4074 TCAGTGTGGTGCACGAGTGGGTATCATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGA 4133  
DB 2025 TCAGTGTGGTGCACGAGTGGGTATCATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGA 1966  
QY 4134 GAGTTTTCGCCCGGAAGAACTTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATGFG 4193  
DB 1965 GAGTTTTCGCCCGGAAGAACTTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATGTC 1906  
QY 4194 CGCGTATTATCCCGTATTGACCGCGGCGAGAGCAACTCGGTGCGCGCATACACTATTC 4253  
DB 1905 TACACTATTATCCCGTATTGACCGCGGCGAGAGCAACTCGGTGCGCGCGGTATTC 1846  
QY 4254 TCAGATGACTTGGTTGAGTACTCACAGTCAACAGAAAGCATCTTACGGATGGCATGAC 4313  
DB 1845 TCAGATGACTTGGTTGAGTACTCACAGTCAACAGAAAGCATCTTACGGATGGCATGAC 1786  
QY 4314 AGTAAGAGAAATTTGCACTGCTGCCATTAACCATGATGATTAACACTGCGGCCAATTACT 4373  
DB 1785 AGTAAGAGAAATTTGCACTGCTGCCATTAACCATGATGATTAACACTGCGGCCAATTACT 1726  
QY 4374 TCTGACAAACGATCGAGAGACCGAAGAGCTAACCGCTTTTTCGACAAACATGGGGGATCA 4433  
DB 1725 TCTGACAAACGATCGAGAGACCGAAGAGCTAACCGCTTTTTCGACAAACATGGGGGATCA 1666  
QY 4434 TGTAACCTCGCTTGCATCGTGGGACCGAGCTGAATGAAGCCATACCAACAGAGAGCG 4493  
DB 1665 TGTAACCTCGCTTGCATCGTGGGACCGAGCTGAATGAAGCCATACCAACAGAGAGCG 1606  
QY 4494 TGACACACGATGCTGTAGCAATGGCAACAACTGTGGCAAACTATTAACTGGCGAAT 4553  
DB 1605 TGACACACGATGCTGTAGCAATGGCAACAACTGTGGCGAACTATTAACTGGCGAAT 1546  
QY 4554 ACTTACTAGCTTCCCGGCAACAAATTAATAGATCGATGAGGCGGATTAAGTTGACG 4613  
DB 1545 ACTTACTAGCTTCCCGGCAACAAATTAATAGATCGATGAGGCGGATTAAGTTGACG 1486  
QY 4614 ACCACTTCTGGCTCGGCTTCCCGGCTTCCCGGCTTCCCGGCTTCCCGGCTTCCCGGCT 4673  
DB 1485 ACCACTTCTGGCTCGGCTTCCCGGCTTCCCGGCTTCCCGGCTTCCCGGCTTCCCGGCT 1426  
QY 4674 TGACGCTGGGTCTCGCGTATCATGTGACGACTGGGCGCAGATGGTAAGCCCTCCCGTAT 4733  
DB 1425 TGACGCTGGGTCTCGCGTATCATGTGACGACTGGGCGCAGATGGTAAGCCCTCCCGTAT 1366  
QY 4734 CGTAGTTATCTACAGGACGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGC 4793  
DB 1365 CGTAGTTATCTACAGGACGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGC 1306  
QY 4794 TGAGATAGTGCCTCACTGATTAAGCATTTGTAAGTGTGACCAAGTTTACTCATATAT 4853  
DB 1305 TGAGATAGTGCCTCACTGATTAAGCATTTGTAAGTGTGACCAAGTTTACTCATATAT 1246  
QY 4854 ACTTTAGATTGAATTAAGTCAATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTT 4913  
DB 1245 ACTTTAGATTGAATTAAGTCAATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTT 1186  
QY 4914 TGATAATCTCATGACCAAAATCCCTTACGTGAGTTTTCGTTCCACTGAGGCTCAGACCC 4973  
DB 1185 TGATAATCTCATGACCAAAATCCCTTACGTGAGTTTTCGTTCCACTGAGGCTCAGACCC 1126  
QY 4974 CGTAGAAAGATCAAGGATCTCTTCAGATCTTTTTCGCGCGTAATCTGCTGCTT 5033  
DB 1125 CGTAGAAAGATCAAGGATCTCTTCAGATCTTTTTCGCGCGTAATCTGCTGCTT 1066  
QY 5034 GCAAAACAAAAACACCGCTTACAGCGGTGTTTGTGTTCCGGATCAAGAGCTACCAAC 5093  
DB 1065 GCAAAACAAAAACACCGCTTACAGCGGTGTTTGTGTTCCGGATCAAGAGCTACCAAC 1006  
QY 5094 TCTTTTCCGAAAGTAACTGGCTTACAGAGCGGAGATACCAAACTACTGTCTCTTCTAGT 5153

DB 1005 TCTTTTCCGAAAGTAACTGGCTTACAGAGCGGAGATACCAAACTACTGTCTCTTCTAGT 946  
QY 5154 GTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT 5213  
DB 945 GTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT 886  
QY 5214 GCTAATCTCTGTACAGTGGCTGCTGCCAGTGGCGATTAAGTGTCTTACCGGGTTGA 5273  
DB 885 GCTAATCTCTGTACAGTGGCTGCTGCCAGTGGCGATTAAGTGTCTTACCGGGTTGA 826  
QY 5274 CTCAGAGCATAGTTACCGGATAAGGCGCACGCTGCGGCTGAAACGGGGGTTCTGTCAC 5333  
DB 825 CTCAGAGCATAGTTACCGGATAAGGCGCACGCTGCGGCTGAAACGGGGGTTCTGTCAC 766  
QY 5334 ACAGCCACGCTTGGAGCGAACCACTACACCGAACTGAGATACCTTACAGCGTGAGCTATG 5393  
DB 765 ACAGCCACGCTTGGAGCGAACCACTACACCGAACTGAGATACCTTACAGCGTGAGCTATG 706  
QY 5394 AGAAAGCGCACTGCTTCCGAAAGGAGAAAGCGGACAGGATTCGGTAAGCGGAGGT 5453  
DB 705 AGAAAGCGCACTGCTTCCGAAAGGAGAAAGCGGACAGGATTCGGTAAGCGGAGGT 646  
QY 5454 CGGAAACGAGAGCGCACGAGGAGCTTCCAGGGGGAACCGCTGCTATCTTTATAGTCC 5513  
DB 645 CGGAAACGAGAGCGCACGAGGAGCTTCCAGGGGGAACCGCTGCTATCTTTATAGTCC 586  
QY 5514 TGTGCGGTTTCCGCACTCTGACTTGAAGCTCGATTTTGTGATGCTCGTCAAGGGGGCG 5573  
DB 585 TGTGCGGTTTCCGCACTCTGACTTGAAGCTCGATTTTGTGATGCTCGTCAAGGGGGCG 526  
QY 5574 GAGCTATGGAAGAAACGCAAGCAACGCGGCTTTTACGGTTCTGCGCTTTTGTCTGCGC 5633  
DB 525 GAGCTATGGAAGAAACGCAAGCAACGCGGCTTTTACGGTTCTGCGCTTTTGTCTGCGC 466  
QY 5634 TTTTGTCTACATGTTCTTCCGCTGTTATCCCTGATTTCTGCGATTAACCGTATTACCGC 5693  
DB 465 TTTTGTCTACATGTTCTTCCGCTGTTATCCCTGATTTCTGCGATTAACCGTATTACCGC 406  
QY 5694 CTTTGTAGTGAAGTCAATCGCTCGCGCAGCGCAACGACCGAGCGCAGCGAGTCAAGTGA 5753  
DB 405 CTTTGTAGTGAAGTCAATCGCTCGCGCAGCGCAACGACCGAGCGCAGCGAGTCAAGTGA 346  
QY 5754 CGAGAAAGCGAAGAGCGCCCAATACGCAAAACCGCTCTCCCGCGCTTGGCCGATCA 5813  
DB 345 CGAGAAAGCGAAGAGCGCCCAATACGCAAAACCGCTCTCCCGCGCTTGGCCGATCA 286  
QY 5814 TTAATGCAG 5822  
DB 285 TTAATGCAG 277

RESULT 11  
LOCUS AV122058 3551 bp DNA circular SVN 28-JUL-2002  
DEFINITION Expression vector pQM411, complete sequence.  
ACCESSION AV122058  
VERSION AV122058.1 GI:22001012  
KEYWORDS  
SOURCE Expression vector pQM411  
ORGANISM Expression vector pQM411  
REFERENCE 1 (bases 1 to 3551)  
AUTHORS Hardy, C.M.  
TITLE Direct Submission  
JOURNAL Submitted (14-JUN-2002) Wildlife, Pests and Diseases, CSIRO  
Sustainable Ecosystems, GPO Box 284, Canberra, ACT 2601, Australia  
FEATURES  
source location/Qualifiers  
1..3551  
/organism="Expression vector pQM411"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:202450"  
66..425  
promoter

```

misc_feature
terminator
CDS
/note="from Human cytomegalovirus immediate early 1 (IE-1)
Gene"
437..456
/note="multiple cloning site"
457..583
/note="Simian virus 40 derived polyA terminator"
1465..2325
/function="ampicillin resistance"
/codon_start=1
/product="beta lactamase"
/protein_id="AAM82245.1"
/db_xref="GI:22001013"
/translation="MSIQHFRVALIPFFAAFLPVPFAHPETLVKVKDAADQLGARVGY
I ELDNSKKILEPERFERRFMMVTFSKYLICHLLIRISIDAGQEQLRRARYSQNDLVE
YSPVTEKLTGTMTVRELCSAAITMTSDNTAANLLLTITGGPKELTAFLEHMGDHTPL
DRWPELNEAI PNDRDPTMTVPAMPTTLTKLLTGELTLASRQOLI EMEADKVKAGPL
LRSLALPWGFTADKSGAGERSGRI IAAIGDPDGKPSRIWVITYTGSQA TMENRQIAT
EIGASLI KHW"

```

## ORIGIN

Query Match	46.1%	Score 2683.4	DB 12	Length 3551
Best Local Similarity	99.4%	Prod. No. 0		
Matches 2693	Conservative 0	Mismatches 16	Indels 0	Gaps 0
Qy	3114	CGGCGGATGCGAGTGGGGCCCAATTGCGCCTATAGTGCCTGTATTAATCAATCACTGG	3173	
Db	612	CGGGAGCATGCGAGTGGGGCCCAATTGCGCCTATAGTGCCTGTATTAATCAATCACTGG	671	
Qy	3174	CGGTGCTTTTACAACGTGTCACATGGGGAACCCCTGCGTTTACCCCACTTAATCGCCCTG	3233	
Db	672	CCGTGCTTTTACAACGTGTCACATGGGGAACCCCTGCGTTTACCCCACTTAATCGCCCTG	731	
Qy	3234	CAGCACATCCCCCTTTTCCGACGTGGCGGTAATAGCGAAGAGGCCCGCACCGATCGCCCTT	3293	
Db	732	CAGCACATCCCCCTTTTCCGACGTGGCGGTAATAGCGAAGAGGCCCGCACCGATCGCCCTT	791	
Qy	3294	CCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGGCG	3353	
Db	792	CCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGGCG	851	
Qy	3354	GGCGGTGTGTGTTACGCGAGGCTGACCGCTACACTTCCGACGCGCCCTAGCGCCGCG	3413	
Db	852	GGCGGTGTGTGTTACGCGAGGCTGACCGCTACACTTCCGACGCGCCCTAGCGCCGCG	911	
Qy	3414	TCCTTTTCGCTTTCTTCCCTTCTCTGCGCACGTTGCGCGGCTTTCGCCGTCAAGCTCT	3473	
Db	912	TCCTTTTCGCTTTCTTCCCTTCTCTGCGCACGTTGCGCGGCTTTCGCCGTCAAGCTCT	971	
Qy	3474	AAATCGGGGGCTCCCTTTAGGGTTCCGAATTTAGAGCTTTACGGCACCTCGACCGCAAAA	3533	
Db	972	AAATCGGGGGCTCCCTTTAGGGTTCCGAATTTAGAGCTTTACGGCACCTCGACCGCAAAA	1031	
Qy	3534	ACTTGAATTTGGGTGATGTTTCAACGTAGTGGGCCATTCGCCCTGTATAGAGCGTTTTTCGCC	3593	
Db	1032	ACTTGAATTTGGGTGATGTTTCAACGTAGTGGGCCATTCGCCCTGTATAGAGCGTTTTTCGCC	1091	
Qy	3594	TTTGACGTGGAGTCCAAGTCTTTTAAATAGTGGACTCTTGTTTCAAACTGGAAACAACCT	3653	
Db	1092	TTTGACGTGGAGTCCAAGTCTTTTAAATAGTGGACTCTTGTTTCAAACTGGAAACAACCT	1151	
Qy	3654	CAACCCCTATCTCGGTCTAATCTTTTGAATTAATAGGGAATTTTCCGATTTTCGGCTATTG	3713	
Db	1152	CAACCCCTATCTCGGTCTAATCTTTTGAATTAATAGGGAATTTTCCGATTTTCGGCTATTG	1211	
Qy	3714	GTTTAAAAATGAGCTGATTTTAAACAAATATTTTAAACGGAATTTTAAACAAATATTAAGCTT	3773	
Db	1212	GTTTAAAAATGAGCTGATTTTAAACAAATATTTTAAACGGAATTTTAAACAAATATTAAGCTT	1271	
Qy	3774	TACAATTTTCGCTGATCGGTATTTTCTTCCCTTACGCACTCTGTCGGGTATTTTCAACCGCA	3833	
Db	1272	TACAATTTTCGCTGATCGGTATTTTCTTCCCTTACGCACTCTGTCGGGTATTTTCAACCGCA	1331	
Qy	3834	TAACAGGTGGCACTTTTTCGGGGAAATGTGCGGGGAACCCCTATTTGTTTATTTTCTTAAT	3893	

• 3834. TACAGGTGGCACTTTTCGGGAAATGTGCGCGGAACCCCTATTGTTTATTTTCTAAAT 3893

Db	1332	 TACAGGTGGCACATTTTCGGGGAAATGTGCGCGAACCCTATTATTGTTATATTTTCTTAANAT	1391
Qy	3894	ACATTCAAAATATGTATCCGCTCATGTAGACAAATAACCCCTGATAAATGCTTCAATAATATG	3953
Db	1392	ACATTCAAAATATGTATCCGCTCATGTAGACAAATAACCCCTGATAAATGCTTCAATAATATG	1451
Qy	3954	AAAAAGGAAGATATGAGTATTCAAACATTTCCGTGTGCCCTTATTTCCCTTTTTTGGCGC	4013
Db	1452	AAAAAGGAAGATATGAGTATTCAAACATTTCCGTGTGCCCTTATTTCCCTTTTTTGGCGC	1511
Qy	4014	ATTTTGCCTTCCTGTTTTTTCCTCACCACAGAAACGCTGTGTGAAAAGTAAAGATGCTGAAGA	4073
Db	1512	ATTTTGCCTTCCTGTTTTTTCCTCACCACAGAAACGCTGTGTGAAAAGTAAAGATGCTGAAGA	1571
Qy	4074	TCAGTTGGGTGACAGAGTGGGTTACATTCGAACTGGATCTCAAACAGCGGTAAAGATCTCTTGA	4133
Db	1572	TCAGTTGGGTGACAGAGTGGGTTACATTCGAACTGGATCTCAAACAGCGGTAAAGATCTCTTGA	1631
Qy	4134	GAGTTTTGCGCCCGGAAGAACGTTTTTCCAAATGATGAGCACATTTTAAAGTTCCTGCTATGTGG	4193
Db	1632	GAGTTTTGCGCCCGGAAGAACGTTTTTCCAAATGATGAGCACATTTTAAAGTTCCTGCTATGTGA	1691
Qy	4194	CGCGGTATTATCCGGTATTGACCGCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTC	4253
Db	1692	TACACTATTATCCCGTATTGACCGCGGGCAAGAGCAACTCGGTGCGCGCGCGGTATTC	1751
Qy	4254	TCAGATGACCTTGGTTGAGTACTCACCAAGTCACAGAAAGCATCTTACGGATGSCATGAC	4313
Db	1752	TCAGATGACCTTGGTTGAGTACTCACCAAGTCACAGAAAGCATCTTACGGATGSCATGAC	1811
Qy	4314	AGTAAGAGAAATTATGCAAGTCTGCCATAACCATGAGTGATAACACTGCGGCCCAACTTACT	4373
Db	1812	AGTAAGAGAAATTATGCAAGTCTGCCATAACCATGAGTGATAACACTGCGGCCCAACTTACT	1871
Qy	4374	TCTGACAAACGATCGGAGGACCGAAGGAGCTAAACCGCTTTTTTTCGCAACAACATGGGGATCA	4433
Db	1872	TCTGACAAACGATCGGAGGACCGAAGGAGCTAAACCGCTTTTTTTCGCAACAACATGGGGATCA	1931
Qy	4434	TGTAACCTCGCTCTGATCGTTGGGAACCGGAGCTGAAATGAAGCCCATACCAACGACGAGCG	4493
Db	1932	TGTAACCTCGCTCTGATCGTTGGGAACCGGAGCTGAAATGAAGCCCATACCAACGACGAGCG	1991
Qy	4494	TGCACCAACGATGCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACT	4553
Db	1992	TGCACCAACGATGCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACT	2051
Qy	4554	ACTTACTCTAGCTTCCGGCAACAAATTAATAGACTGATGGAGCGGATTAAGTTGCGAGG	4613
Db	2052	ACTTACTCTAGCTTCCGGCAACAAATTAATAGACTGATGGAGCGGATTAAGTTGCGAGG	2111
Qy	4614	ACCACTTCTGCGCTCGGCCCTTCGGCTGCGCTGGTTTATTGCTGATAAAATCTCGAGCCGG	4673
Db	2112	ACCACTTCTGCGCTCGGCCCTTCGGCTGCGCTGGTTTATTGCTGATAAAATCTCGAGCCGG	2171
Qy	4674	TGAGCTGGGTCTCGGGTATCAATTCAGACACTGGGGCCAGATGGTTAAGCCCTCCCGTAT	4733
Db	2172	TGAGCTGGGTCTCGGGTATCAATTCAGACACTGGGGCCAGATGGTTAAGCCCTCCCGTAT	2231
Qy	4734	CGTAGTTTATCTACACGACGGGGAGTCAGGCACCTATGGATGCAACGAATAGACAGATCGC	4793
Db	2232	CGTAGTTTATCTACACGACGGGGAGTCAGGCACCTATGGATGCAACGAATAGACAGATCGC	2291
Qy	4794	TGAGATAGGTGCCCTCACTGATTAAGCAATTGGTAACCTGTAGACCAAGTCTTACTCATATAT	4853
Db	2292	TGAGATAGGTGCCCTCACTGATTAAGCAATTGGTAACCTGTAGACCAAGTCTTACTCATATAT	2351
Qy	4854	ACTTTAGATTGATTTAAAACTTCAATTTTTTAAATGAGGATCTAGGTGAAGATCTCTTTT	4913
Db	2352	ACTTTAGATTGATTTAAAACTTCAATTTTTTAAATGAGGATCTAGGTGAAGATCTCTTTT	2411
Qy	4914	TGATAATCTCATGACCAAAATCCCTTAAACGTGAGTTTTTCGTTCCACTGACGGTCAAGCCC	4973



|||||  
1312 CAACCCCTATCTCGGTCTATCTCTTTTGAATTAAGGGATTTTGGCGATTTTCGGCTATTG 1371  
QY  
3714 GTTAAAAAATGAGCTGATTTTAAACAATAATTTAAACCGAAATTTTAAACAAAATATTAAACGTT 3773  
Db  
1372 GTTAAAAAATGAGCTGATTTTAAACAATAATTTAAACCGAAATTTTAAACAAAATATTAAACGTT 1431  
QY  
3774 TACAATTTTCGCTGATGCGGTATTTTCTCTTAACGCAATCTGTGCGGTATTTTCAACCGCA 3833  
Db  
1432 TACAATTTTCGCTGATGCGGTATTTTCTCTTAACGCAATCTGTGCGGTATTTTCAACCGCA 1491  
QY  
3834 TACAGGTGGCACTTTTCGGGGAATGTGGCGGAACCCCTATTTTCTTAAT 3893  
Db  
1492 TACAGGTGGCACTTTTCGGGGAATGTGGCGGAACCCCTATTTTCTTAAT 1551  
QY  
3894 ACATTTCAAAATATGATTCGCTCATGAGCAATAACCCCTGATAATGCTTCAATATATG 3953  
Db  
1552 ACATTTCAAAATATGATTCGCTCATGAGCAATAACCCCTGATAATGCTTCAATATATG 1611  
QY  
3954 AAAAGGAAGATGATGATTTCAACATTTTCGCTGTGCGCTTATTCCTTTTTCGGC 4013  
Db  
1612 AAAAGGAAGATGATGATTTCAACATTTTCGCTGTGCGCTTATTCCTTTTTCGGC 1671  
QY  
4014 ATTTTCCTCTCTGTTTCTGCTACCCAGAAACGCTGTGGAAGTAAAGATGCTGAAGA 4073  
Db  
1672 ATTTTCCTCTCTGTTTCTGCTACCCAGAAACGCTGTGGAAGTAAAGATGCTGAAGA 1731  
QY  
4074 TCAGTTGGGTGCAGAGTGGGTACATCGAATCTGATCTCAACAGCGGTAAAGATGCTTGA 4133  
Db  
1732 TCAGTTGGGTGCAGAGTGGGTACATCGAATCTGATCTCAACAGCGGTAAAGATGCTTGA 1791  
QY  
4134 GAGTTTTCGCCCCGAAAGAGTTTTCATGATGAGCACTTTTAAAGTTCGTCTATG 4193  
Db  
1792 GAGTTTTCGCCCCGAAAGAGTTTTCATGATGAGCACTTTTAAAGTTCGTCTATG 1851  
QY  
4194 CGGGTATTAATCCGTTATTCAGCGCGGCAAGAGCACTCGGTCCGCGATACACTATTC 4253  
Db  
1852 TACACTATTAATCCGTTATTCAGCGCGGCAAGAGCACTCGGTCCGCGCGGATTC 1911  
QY  
4254 TCAGAAATGACTTGGTTGAGTACTCAACAGTACAGAAAGATCTTACGGATGGCATGAC 4313  
Db  
1912 TCAGAAATGACTTGGTTGAGTACTCAACAGTACAGAAAGATCTTACGGATGGCATGAC 1971  
QY  
4314 AGTAAGAAATTAATGCAATGCTGCTCAATACATGAGTGAATCACTGCGGCCAACTTACT 4373  
Db  
1972 AGTAAGAAATTAATGCAATGCTGCTCAATACATGAGTGAATCACTGCGGCCAACTTACT 2031  
QY  
4374 TCTGACAAACGATCGGAGGACCGAGGAGCTAAACGCTTTTTCACAAACATGGGGATCA 4433  
Db  
2032 TCTGACAAACGATCGGAGGACCGAGGAGCTAAACGCTTTTTCACAAACATGGGGATCA 2091  
QY  
4434 TGTAACCTCGCTTGTATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCG 4493  
Db  
2092 TGTAACCTCGCTTGTATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCG 2151  
QY  
4494 TGACACCAAGATGCTGATGCAATGGAACAACGTTGCGCAAACTATTAATCTGGCGAAT 4553  
Db  
2152 TGACACCAAGATGCTGATGCAATGGAACAACGTTGCGCAAACTATTAATCTGGCGAAT 2211  
QY  
4554 ACTTACTCTAGCTTCCCGGCAACAAATTAATAGACTGGAAGCGGATGAAGTTCGAGG 4613  
Db  
2212 ACTTACTCTAGCTTCCCGGCAACAAATTAATAGACTGGAAGCGGATGAAGTTCGAGG 2271  
QY  
4614 ACCACTTCTCGCTCGGCCCTCTCGGCTGGCTGTTTATGCTGATAAATCTGAGCCGG 4673  
Db  
2272 ACCACTTCTCGCTCGGCCCTCTCGGCTGGCTGTTTATGCTGATAAATCTGAGCCGG 2331  
QY  
4674 TGAGCGTGGTCTCGCGGTATCAATGAGCACTGAGGAGGAGTGGTAAAGCCCTCCCGTAT 4733  
Db  
2332 TGAGCGTGGTCTCGCGGTATCAATGAGCACTGAGGAGGAGTGGTAAAGCCCTCCCGTAT 2391  
QY  
4734 CGTAGTTATCTACAGCGGGAGTCAAGGCACTATGATGAACGAAATAGACAGATCGC 4793

2392 CGTAGTTATCTACAGCGGGAGTCAGCGCAACTATGATGAACGAATAGACAGATCGC 2451  
QY  
4794 TGAGATAGTGGCTCTCACTGATTAAGCAATGCTACTGTCAGACCAAGTTTACTCATATAT 4853  
Db  
2452 TGAGATAGTGGCTCTCACTGATTAAGCAATGCTACTGTCAGACCAAGTTTACTCATATAT 2511  
QY  
4854 ACTTTAGATTAATTTAAACCTTCAATTTTAAATTTAAAGGATCTAGGTGAAGATCTCTTTT 4913  
Db  
2512 ACTTTAGATTAATTTAAACCTTCAATTTTAAATTTAAAGGATCTAGGTGAAGATCTCTTTT 2571  
QY  
4914 TGATAATCTCATGACCAAAATCCCTTAACGCTGAGTTTTCGTTCCACTGAGCGTCAGACCC 4973  
Db  
2572 TGATAATCTCATGACCAAAATCCCTTAACGCTGAGTTTTCGTTCCACTGAGCGTCAGACCC 2631  
QY  
4974 CGTAGAAAAGATCAAAAGGATCTTCTTGAGATCTCTTTTTCGCGCGTAATCTGCTGCTT 5033  
Db  
2632 CGTAGAAAAGATCAAAAGGATCTTCTTGAGATCTCTTTTTCGCGCGTAATCTGCTGCTT 2691  
QY  
5034 GCMAAAGAAAACCAACCGCTACAGCGGTGTTTGTTCGCGGATCAAGAGCTTACCAAC 5093  
Db  
2692 GCMAAAGAAAACCAACCGCTACAGCGGTGTTTGTTCGCGGATCAAGAGCTTACCAAC 2751  
QY  
5094 TCTTTTTCGGAAGTAACTGCTTTCAGCAGAGCGCAGATACCAATACCTCTCTAGT 5153  
Db  
2752 TCTTTTTCGGAAGTAACTGCTTTCAGCAGAGCGCAGATACCAATACCTCTCTAGT 2811  
QY  
5154 GTAGCCGTAGTTAGGCCCAACCTTCAAGAACTCTGTAGCAGCCGCTACATACCTCGCTCT 5213  
Db  
2812 GTAGCCGTAGTTAGGCCCAACCTTCAAGAACTCTGTAGCAGCCGCTACATACCTCGCTCT 2871  
QY  
5214 GCTAATCTGTTACAGTGGCTGCTGCAAGTGGCGATAAGTCTGCTTACCGGGTTGGA 5273  
Db  
2872 GCTAATCTGTTACAGTGGCTGCTGCAAGTGGCGATAAGTCTGCTTACCGGGTTGGA 2931  
QY  
5274 CTCAGAAGCATAGTTACCGGATAAGCGCAGCGCTCGGCTGAACGGGGGTTCGTGAC 5333  
Db  
2932 CTCAGAAGCATAGTTACCGGATAAGCGCAGCGCTCGGCTGAACGGGGGTTCGTGAC 2991  
QY  
5334 ACAGCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAAGTATG 5393  
Db  
2992 ACAGCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAAGTATG 3051  
QY  
5394 AGAAGGCGACGCTTCCGAAAGGAGAAAGCGGACAGGTATCCGGTAAAGCGCAGGCT 5453  
Db  
3052 AGAAGGCGACGCTTCCGAAAGGAGAAAGCGGACAGGTATCCGGTAAAGCGCAGGCT 3111  
QY  
5454 CGCAACAGGAGCGCACAGGAGGCTTCCAGGGGAAACGCTGCTATCTTTATAGTCC 5513  
Db  
3112 CGCAACAGGAGCGCACAGGAGGCTTCCAGGGGAAACGCTGCTATCTTTATAGTCC 3171  
QY  
5514 TGTGGGTTTCGCCACCTCTGACCTTGAAGCTTCGATTTTGTGATGCTCGTCAAGGGGGG 5573  
Db  
3172 TGTGGGTTTCGCCACCTCTGACCTTGAAGCTTCGATTTTGTGATGCTCGTCAAGGGGGG 3231  
QY  
5574 GAGCCTATGGAAGAAACGCGAGCAACGGGGCTTTTACGGTCTCTGCGCTTTTCTGCGCC 5633  
Db  
3232 GAGCCTATGGAAGAAACGCGAGCAACGGGGCTTTTACGGTCTCTGCGCTTTTCTGCGCC 3291  
QY  
5634 TTTTGTCTCATATCTCTTCTGCTGTTATCCCTGATTTCTGTGATAACCGTATTTACCGC 5693  
Db  
3292 TTTTGTCTCATATCTCTTCTGCTGTTATCCCTGATTTCTGTGATAACCGTATTTACCGC 3351  
QY  
5694 CTTTGTGTGAGCTGATACCGCTCGCGCAGCGAAACGACCGAGCGCAGCGAGTCAAGTGA 5753  
Db  
3352 CTTTGTGTGAGCTGATACCGCTCGCGCAGCGAAACGACCGAGCGCAGCGAGTCAAGTGA 3411  
QY  
5754 CGAGGAAGCGGAAGAGCGCCCAATACGCAACCGCTCTCCCGCGGTTGGCGGATCA 5813  
Db  
3412 CGAGGAAGCGGAAGAGCGCCCAATACGCAACCGCTCTCCCGCGGTTGGCGGATCA 3471  
QY  
5814 TTAATTCAG 5822  
Db  
3472 TTAATTCAG 3480

RESULT 13  
LOCUS CVGEM7LICF/c 3033 bp DNA circular SYN 16-MAY-1995  
DEFINITION Ligation-independent cloning vector pGEM-7Zf(+)/LIC-F, complete sequence.  
ACCESSION U25272.1 GI:806899  
KEYWORDS beta-lactamase; ligation-independent cloning region; Nari restriction site.  
SOURCE Cloning vector pGEM-7Zf(+)/LIC-F  
ORGANISM artificial sequences; vectors.  
REFERENCE 1 (bases 1 to 3033)  
AUTHORS Haun, R.S., Serventi, I.M. and Moss, J.  
TITLE Rapid, reliable ligation-independent cloning of PCR products using modified plasmid vectors  
JOURNAL Biotechniques 13 (4), 515-518 (1992)  
MEDLINE 93119606  
PUBMED 1362067  
REFERENCE 2 (bases 1 to 3033)  
AUTHORS Haun, R.S.  
JOURNAL Direct Submission  
SUBMITTED (19-APR-1995) Randy S. Haun, NIH, NHLBI, Laboratory of Cellular Metabolism, Bldg. 10, Room 5N307, Bethesda, MD 20892-1434, USA  
COMMENT Ligation-independent cloning vector. Preparation of the vector for cloning includes linearization with Nari, gel purification of the linearized vector, and treatment with T4 DNA polymerase in the presence of dATP. The target sequence can be amplified using sequence specific primers modified at the 5' end to contain an additional 13 nucleotides complementary to the vector. The forward primer should contain 5'-CTGGTTCGCGA-3', followed by 12-15 nucleotide target-specific sequence. The reverse primer should contain 5'-CTCGCTCGCGA-3' followed by 12-15 nucleotide target-specific sequence. The amplified sequence should be gel purified and treated with T4 DNA polymerase in the presence of dTTP. Annealing of the vector and the amplification product forms a duplex that can be used directly for transformation. Sequences amplified using these primers are also compatible with the pBluescript II KS(+)/LIC vector (ATCC 87047) and pGEM-7Zf(+)/LIC-R vector (ATCC 87049). Differs from pGEM-7Zf(+)/LIC-R (ATCC 87049) only in the orientation of complementary ends generated at the cloning site. This vector has been deposited in the ATCC repository (ATCC 87048).

FEATURES  
source Location/Qualifiers  
1..3033  
/organism="Cloning vector pGEM-7Zf(+)/LIC-F"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:39798"  
/note="pGEM-7Zf(+)/LIC-F; ATCC 87048"  
promoter join(3017..3033,1..6)  
/note="bacteriophage T7 promoter"  
misc\_feature 55..81  
/product="ligation-independent cloning region"  
/evidence=experimental  
misc\_feature 66..71  
/product="Nari restriction site"  
/evidence=experimental  
promoter complement(151..173)  
/note="bacteriophage SP6 promoter"  
gene complement(193..195)  
/gene="lacZ"  
CDS complement(<193..195)  
/gene="lac2"  
/note="initiation codon"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="AAA66386.1"  
/db\_xref="GI:806890"  
/translation="M"  
CDS complement(1352..2212)

```
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AAA66387.1"
/db_xref="GI:806891"
/translation="MSIQHFRVALIPPPAAFLPVPAPHPETLVKVKDAEDQLGARVGY
IQLDLSNGIKILDSFRPERRPMSTFKVLLCGAVLSRIDAGQBOGLRRIRHYGNDLVE
YSPVTERLITGDGMTVRELCSAAITMSDNTANLLITGGPKELTAFLHMGDGHVTRL
DRWPELAEAIPIIDERDTTPVAMATTLRLKLLTGLLELLTSLASQQILDMNEADKVGPL
LRSLALPAGWFIADKSGAGERSGIIAALGPDGKPSRIWVIVTTGSAQTMDERNRQIA
EIGASLIKHW"
complement(2398..2853)
/standard_name="f1 origin"

rep_origin
ORIGIN
Query Match      46.0%; Score 2677; DB 12; Length 3033;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3146 TATAGTGAGTCGTATTACAATTCACTGGCGGCGTGTGTTTACAACGTCGTGACGGGAAAC 3205
DB 3033 TATAGTGAGTCGTATTACAATTCACTGGCGGCGTGTGTTTACAACGTCGTGACGGGAAAC 2974

QY 3206 CTTGGCGGTACCCAACTTAATGCGCTTCGACGACATCCCTTTCCGACAGCTGGCGTAAT 3265
DB 2973 CTTGGCGGTACCCAACTTAATGCGCTTCGACGACATCCCTTTCCGACAGCTGGCGTAAT 2914

QY 3266 AGCGAAGAGGCGCGCACCGATCGCCCTTCCCAACAGTTGCGCAGGCTGAATGGCGAATGG 3325
DB 2913 AGCGAAGAGGCGCGCACCGATCGCCCTTCCCAACAGTTGCGCAGGCTGAATGGCGAATGG 2854

QY 3326 ACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGTGTGGTGTGTTACGCGCAGCGTGACCG 3385
DB 2853 ACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGTGTGGTGTGTTACGCGCAGCGTGACCG 2794

QY 3386 CTACACTGCGCAGCGCGCTAGCGCGCGCTCTTCGCTTTCTTCCCTTCTTCTTCGCGCA 3445
DB 2793 CTACACTGCGCAGCGCGCTAGCGCGCGCTCTTCTTCCCTTCTTCTTCGCGCA 2734

QY 3446 CGTTTCGCGCGCTTTCCCGCTCAAGCTCTAAATCGGGGCGCTCCCTTTAGGGTTCCGATTTA 3505
DB 2733 CGTTTCGCGCGCTTTCCCGCTCAAGCTCTAAATCGGGGCGCTCCCTTTAGGGTTCCGATTTA 2674

QY 3506 GAGCTTTACGGCAGCTCGACCGCAAAACTTGATTTGGGTGATGCTTCACTAGTGGCG 3565
DB 2673 GAGCTTTACGGCAGCTCGACCGCAAAACTTGATTTGGGTGATGCTTCACTAGTGGCG 2614

QY 3566 CATGCGCCCTGATAGACGGTTTTTCGCCCTTTGAGCTTCGAGTCCACGTTCTTTAATAGTG 3625
DB 2613 CATGCGCCCTGATAGACGGTTTTTCGCCCTTTGAGCTTCGAGTCCACGTTCTTTAATAGTG 2554

QY 3626 GACTCTTTGTTCCAAACTGGAAACAACCTCAACCCCTATCTCGGTCTATTCTTTTGATTTAT 3685
DB 2553 GACTCTTTGTTCCAAACTGGAAACAACCTCAACCCCTATCTCGGTCTATTCTTTTGATTTAT 2494

QY 3686 AAGGATTTTGGCGATTTGCGGCTATTGGTTTAAAAAATGAGCTGATTTTAAACAATTTTA 3745
DB 2493 AAGGATTTTGGCGATTTGCGGCTATTGGTTTAAAAAATGAGCTGATTTTAAACAATTTTA 2434

QY 3746 ACGCGAATTTTAAACAAATATTAACGTTTACAAATTCGCCCTGATGCGGTATTCTTCCTTT 3805
DB 2433 ACGCGAATTTTAAACAAATATTAACGTTTACAAATTCGCCCTGATGCGGTATTCTTCCTTT 2374

QY 3806 ACGCATCTGTCGGTATTTCACCGCATACAGGTGGCACTTTTTCGGGGAATGTCGGCG 3865
DB 2373 ACGCATCTGTCGGTATTTCACCGCATACAGGTGGCACTTTTTCGGGGAATGTCGGCG 2314

QY 3866 GAACCCCTATTGTTTATTTCTTAAATACATTCGAATATGTTATCCGCTCATGACACAAT 3925
DB 2313 GAACCCCTATTGTTTATTTTCTTAAATACATTCGAATATGTTATCCGCTCATGACACAAT 2254

QY 3926 AACCTGATTAATGCTTCAATAATATTGAAAAAGGAGATGATGATGATTCACAAATTTCC 3985
```

Db 2253 AACCTGATTAATGCTTCAATAATAATATTTGAAAAAGAGAGATATGAGTATTCAACATTTC 2194  
Qy 3986 GTGTGCGCCCTATTCCCTTTTTCGGCGCATTTTCCTCTCTGTTTTCCTCAACCGAGAA 4045  
Db 2193 GTGTGCGCCCTATTCCCTTTTTCGGCGCATTTTCCTCTCTGTTTTCCTCAACCGAGAA 2134  
Qy 4046 CGCTGTGTAAGTAAAGATGCTGAAGATCAGTTGGGTGACAGTGGGTTCATCATGAAC 4105  
Db 2133 CGCTGTGTAAGTAAAGATGCTGAAGATCAGTTGGGTGACAGTGGGTTCATCATGAAC 2074  
Qy 4106 TGGATCTCAACAGCGGTAAGATCCTTCAGAGTATTTTCGCGCGGAGAGAGTTCCTCAATGA 4165  
Db 2073 TGGATCTCAACAGCGGTAAGATCCTTCAGAGTATTTTCGCGCGGAGAGAGTTCCTCAATGA 2014  
Qy 4166 TGGATCTCAACAGCGGTAAGATCCTTCAGAGTATTTTCGCGCGGAGAGAGTTCCTCAATGA 4165  
Db 2013 TGGATCTCAACAGCGGTAAGATCCTTCAGAGTATTTTCGCGCGGAGAGAGTTCCTCAATGA 4225  
Qy 4225 TGGATCTCAACAGCGGTAAGATCCTTCAGAGTATTTTCGCGCGGAGAGAGTTCCTCAATGA 4225  
Db 1954 TGGATCTCAACAGCGGTAAGATCCTTCAGAGTATTTTCGCGCGGAGAGAGTTCCTCAATGA 1954  
Qy 4226 AGCAACTCGGTGCGCGCATACACTATTCTCAGATGACTTGGTGTGAGTACTCAGCAGTCA 4285  
Db 1953 AGCAACTCGGTGCGCGCATACACTATTCTCAGATGACTTGGTGTGAGTACTCAGCAGTCA 1894  
Qy 4286 CAGAAAGCATCTTACCGATGGCATGACAGTAAAGAGATTTATGAGTGTGCCATACCA 4345  
Db 1893 CAGAAAGCATCTTACCGATGGCATGACAGTAAAGAGATTTATGAGTGTGCCATACCA 1834  
Qy 4346 TGGATGATAACACTGCGCGCAACTTACTTCTGCAACAGATCGGAGACCGAGGAGCTAA 4405  
Db 1833 TGGATGATAACACTGCGCGCAACTTACTTCTGCAACAGATCGGAGACCGAGGAGCTAA 1774  
Qy 4405 CCGCTTTTTCGCAACATCGGGGATCATGTAATCTGCAACAGATCGGAGACCGAGGAGCTAA 4465  
Db 1773 CCGCTTTTTCGCAACATCGGGGATCATGTAATCTGCAACAGATCGGAGACCGAGGAGCTAA 1714  
Qy 4466 TGAATGAAGCATACCAACAGAGAGCGTACACAGATGCTGAGCAATGCAACAA 4525  
Db 1713 TGAATGAAGCATACCAACAGAGAGCGTACACAGATGCTGAGCAATGCAACAA 1654  
Qy 4526 CGTTGCGCAAACTATTAACCTGGGAACTACTTACTCTAGCTTCCCGGCAACAAATTAATAG 4585  
Db 1653 CGTTGCGCAAACTATTAACCTGGGAACTACTTACTCTAGCTTCCCGGCAACAAATTAATAG 1594  
Qy 4586 ACTGATGAGAGCGGATAAAGTTGCGAGGACACTTCTGCGCTCGGCTTCCCGCTGCGCT 4645  
Db 1593 ACTGATGAGAGCGGATAAAGTTGCGAGGACACTTCTGCGCTCGGCTTCCCGCTGCGCT 1534  
Qy 4646 GGTATTGCTGATTAATCTGAGCGCGTGCAGGCTGCTCGGCTATCATATTGCGAGCAC 4705  
Db 1533 GGTATTGCTGATTAATCTGAGCGCGTGCAGGCTGCTCGGCTATCATATTGCGAGCAC 1474  
Qy 4706 TGGGGCGAGATGGTAAAGCCCTCCCGTATCGTAGTTATCTACAGCGGGAGTCAAGCAA 4765  
Db 1473 TGGGGCGAGATGGTAAAGCCCTCCCGTATCGTAGTTATCTACAGCGGGAGTCAAGCAA 1414  
Qy 4766 CTATGATGAACGGAATGACAGATCGCTGAGATGAGTGTCTCATGATTAAGCATTTGGT 4825  
Db 1413 CTATGATGAACGGAATGACAGATCGCTGAGATGAGTGTCTCATGATTAAGCATTTGGT 1354  
Qy 4826 AACTGTGAGCAAGTTTACTCATATATCTTTAGATTGATTTAAACTTCATTTTAAAT 4885  
Db 1353 AACTGTGAGCAAGTTTACTCATATATCTTTAGATTGATTTAAACTTCATTTTAAAT 1294  
Qy 4886 TTAAGAGATCTAGGTGAAGATCCTTTTGTATATCTCATGACCAAAATCCCTTAACGTG 4945  
Db 1293 TTAAGAGATCTAGGTGAAGATCCTTTTGTATATCTCATGACCAAAATCCCTTAACGTG 1234  
Qy 4946 AGTTTTCGCTGAGCGTCAAGCCCGTAGAAGATCAAGAGATCTTCTTGAGATC 5005  
Db 1233 AGTTTTCGCTGAGCGTCAAGCCCGTAGAAGATCAAGAGATCTTCTTGAGATC 1174  
Qy 5006 CTTTTCGCGGTAAATCTGCTTTGCAACCAAAAAACCAACCCCTACCGCGGTGG 5065  
Db 1173 CTTTTCGCGGTAAATCTGCTTTGCAACCAAAAAACCAACCCCTACCGCGGTGG 1114

Qy 5066 TTTGTTTTCGCGATCAAGAGCTACCAACTCTTTTTCGAAAGGTAACTGGCTTCAGCAGAG 5125  
Db 1113 TTTGTTTTCGCGATCAAGAGCTACCAACTCTTTTTCGAAAGGTAACTGGCTTCAGCAGAG 1054  
Qy 5126 CGCAGATACCAATATCTGCTCTTAGTGTAGCGTAGTTAGGCACACACTTCAAGAACT 5185  
Db 1053 CGCAGATACCAATATCTGCTCTTAGTGTAGCGTAGTTAGGCACACACTTCAAGAACT 994  
Qy 5186 CTGTAGCAGCGCTACATACCTCTGCTCTTAATCTGTTTACAGTGGCTGCTGCCAGTG 5245  
Db 993 CTGTAGCAGCGCTACATACCTCTGCTCTTAATCTGTTTACAGTGGCTGCTGCCAGTG 934  
Qy 5246 GCGATTAAGTCTGCTCTTACCGGGTTGGACTCAAGACGATAGTTTACCGGATAAGCGCAGC 5305  
Db 933 GCGATTAAGTCTGCTCTTACCGGGTTGGACTCAAGACGATAGTTTACCGGATAAGCGCAGC 874  
Qy 5306 GGTGCGGCTGAACCGGGGGTTCGTGACACAGCCAGCTTGGAGCGAAGCATACACCG 5365  
Db 873 GGTGCGGCTGAACCGGGGGTTCGTGACACAGCCAGCTTGGAGCGAAGCATACACCG 814  
Qy 5366 AACTGAGATACCTTACAGCGTGAAGTATGAGAAAGCGCCAGCTTCCGAAAGGAGAAAGG 5425  
Db 813 AACTGAGATACCTTACAGCGTGAAGTATGAGAAAGCGCCAGCTTCCGAAAGGAGAAAGG 754  
Qy 5426 CGGACAGGTATCCGTAAGCGGCGAGGGTCCGAAACAGGAGAGCGCACGAGGAGCTTCCAG 5485  
Db 753 CGGACAGGTATCCGTAAGCGGCGAGGGTCCGAAACAGGAGAGCGCACGAGGAGCTTCCAG 694  
Qy 5486 GGGGAAACGCTGATCTTTTATAGTCTCTGCTGCGGTTTCCGCACTCTGACTTGAAGGTC 5545  
Db 693 GGGGAAACGCTGATCTTTTATAGTCTCTGCTGCGGTTTCCGCACTCTGACTTGAAGGTC 634  
Qy 5546 GATTTTGTGATGCTGCTCAGCGGGCGGAGCGCTTATGAAAAAGCCAGCAACCGCGCT 5605  
Db 633 GATTTTGTGATGCTGCTCAGCGGGCGGAGCGCTTATGAAAAAGCCAGCAACCGCGCT 574  
Qy 5606 TTTTACGCTTCTGCGCTTTTGTGCGCTTTTGTCTCATGTTCTTCTGCTGCTTATCCC 5665  
Db 573 TTTTACGCTTCTGCGCTTTTGTGCGCTTTTGTCTCATGTTCTTCTGCTGCTTATCCC 514  
Qy 5666 CTGATTTCTGAGTAACTGATTTACCGCTTTCAGTGTGAGCTGATACCGCTCGCGCAGCC 5725  
Db 513 CTGATTTCTGAGTAACTGATTTACCGCTTTCAGTGTGAGCTGATACCGCTCGCGCAGCC 454  
Qy 5726 GAACGACCGAGCGAGCGAGTCAAGTGAAGCGGAGCGGAGAGCGCCCAATACGCAAC 5785  
Db 453 GAACGACCGAGCGAGCGAGTCAAGTGAAGCGGAGCGGAGAGCGCCCAATACGCAAC 394  
Qy 5786 CGCTCTCCCGCGCTGCGCTTCCGATTCATTAATGCGAG 5822  
Db 393 CGCTCTCCCGCGCTGCGCTTCCGATTCATTAATGCGAG 357

## RESULT 14

CVGEM7LICR/c  
LOCUS 3033 bp DNA circular SYN 16-MAY-1995  
DEFINITION Ligation-independent cloning vector pGEM-7zf(+)/LIC-R, complete  
sequence.  
ACCESSION U25268  
VERSION U25268.1 GI:806878  
KEYWORDS beta-lactamase; ligation-independent cloning region; NarI  
restriction site.  
SOURCE Cloning vector pGEM-7zf(+)/LIC-R  
ORGANISM Cloning vector pGEM-7zf(+)/LIC-R  
REFERENCE 1 (bases 1 to 3033)  
AUTHORS Haun, R.S., Serventi, I.M. and Moss, J.  
TITLE Rapid, reliable ligation-independent cloning of PCR products using  
modified plasmid vectors  
JOURNAL BioTechniques 13 (4), 515-518 (1992)  
MEDLINE 93119606  
PUBMED 1362067



REFERENCE	2 (bases 1 to 3033)	Db	3033	TATAGTCAGTGTGTTATTACAAATTTACTGGCCGTCGTTTTTACAAACGTCGTCGAGTGGGAAAC	2974
AUTHORS	Haun,R.S.	Qy	3206	CCTGGCGTTACCCAACTTAATCGCCTTTCGACACATCCCTTTTCGCCAGCTGGCGTAAT	3265
TITLE	Direct Submission	Db	2973	CCTGGCGTTACCCAACTTAATCGCCTTTCGACACATCCCTTTTCGCCAGCTGGCGTAAT	2914
JOURNAL	Submitted (19-APR-1995) Randy S. Haun, NIH, NHLBI, Laboratory of Cellular Metabolism, Bldg. 10, Room 5N307, Bethesda, MD 20892-1434, USA	Qy	3266	AGCGAAGAGGCCCGCAGATCGATCGCCCTTCCAAACAGTTGCGCAGCTCGAATGGCGAATGG	3325
COMMENT	Ligation-independent cloning vector. Preparation of the vector for cloning includes linearization with NaeI, gel purification of the linearized vector, and treatment with T4 DNA polymerase in the presence of dATP. The target sequence can be amplified using sequence specific primers modified at the 5' end to contain an additional 13 nucleotides complementary to the vector. The forward primer should contain 5'-CTGTTCCGGCA-3', followed by 12-15 nucleotide target-specific sequence. The reverse primer should contain 5'-CTGCTCCGGCA-3', followed by 12-15 nucleotide target-specific sequence. The amplified sequence should be gel purified and treated with T4 DNA polymerase in the presence of dATP. Annealing of the vector and the amplification product forms a duplex that can be used directly for transformation. Sequences amplified using these primers are also compatible with the pBluescript II KS(+)/LIC vector (ATCC 87047) and pGEM-72f(+)/LIC-P vector (ATCC 87048). Differs from pGEM-72f(+)/LIC-F (ATCC 87048) only in the orientation of complementary ends generated at the cloning site. This vector has been deposited in the ATCC repository (ATCC 87049).	Db	2913	AGCGAAGAGGCCCGCAGATCGCCTTCCAAACAGTTGCGCAGCTCGAATGGCGAATGG	2854
FEATURES	Location/Qualifiers	Qy	3326	ACGCGCCCTCTAGCGGCGCATTAAGCGGCGGGGTGGTGGTAAOGCGCAGCGTGACCG	3385
source	1. .3033	Db	2853	ACGCGCCCTCTAGCGGCGCATTAAGCGGCGGGGTGGTGGTAAOGCGCAGCGTGACCG	2794
	/organism="Cloning vector pGEM-72f<+>/LIC-R"	Qy	3386	CTACACTTGGCAGCGCCCTAGCGCCCGCTCTCTTCGCTTCTTCCCTCTCTTCTTCGCA	3445
	/mol_type="genomic DNA"	Db	2793	CTACACTTGGCAGCGCCCTAGCGCCCGCTCTCTTCGCTTCTTCCCTCTCTTCTTCGCA	2734
	/db_xref="taxon:39794"	Qy	3446	CGTTTCGCGGCTTTCGCCGTCAAGCTCTAAATCGGGGCTCCCTTTAGGGTTCGATTTA	3505
promoter	/notes="pGEM-72f(+)/LIC-R; ATCC 87049"	Db	2733	CGTTTCGCGGCTTTCGCCGTCAAGCTCTAAATCGGGGCTCCCTTTAGGGTTCGATTTA	2674
misc_feature	join(3017..3033,1..6)	Qy	3506	GAGCTTTACGCGACCTCGACCGCAAAACTTGATTTGGGTGATGGTTCACTAGTGGCG	3565
	/notes="bacteriophage T7 promoter"	Db	2673	GAGCTTTACGCGACCTCGACCGCAAAACTTGATTTGGGTGATGGTTCACTAGTGGCG	2614
	55..81	Qy	3566	CATCGCCCTGATAGACGGTTTTCGCCCTTTGAGCTTGGAGTCCACGTTCTTTAATAGTG	3625
	/product="ligation-independent cloning region"	Db	2613	CATCGCCCTGATAGACGGTTTTCGCCCTTTGAGCTTGGAGTCCACGTTCTTTAATAGTG	2554
	/evidence=experimental	Qy	3626	GACTCTTGTTCCTCAAACTGGAACAACTCACTAAACCCCTATCTCGCTCTATCTTTGAT	3685
	65..70	Db	2553	GACTCTTGTTCCTCAAACTGGAACAACTCACTAAACCCCTATCTCGCTCTATCTTTGAT	2494
promoter	/evidence=experimental	Qy	3686	AAGGATTTTCGCGATTTTCGCCCTTATTTGGTTTAAATAATGAGCTGATTAACAATATTA	3745
gene	/product="NaeI restriction site"	Db	2493	AAGGATTTTCGCGATTTTCGCCCTTATTTGGTTTAAATAATGAGCTGATTAACAATATTA	2434
	complement(151..173)	Qy	3746	ACGCGAATTTTAAACAAATATTAACGTTTACAAATTCGCCCTGATGCGGTATTTCTCCTT	3805
	/notes="bacteriophage SP6 promoter"	Db	2433	ACGCGAATTTTAAACAAATATTAACGTTTACAAATTCGCCCTGATGCGGTATTTCTCCTT	2374
	complement(193..195)	Qy	3806	ACGCACTGTGGGTATTTTCAACCGCATACAGGTGGCATCTTTTCGGGGAAATGCGCGG	3865
	/gene="lacZ"	Db	2373	ACGCACTGTGGGTATTTTCAACCGCATACAGGTGGCATCTTTTCGGGGAAATGCGCGG	2314
	/complement(<193..195)	Qy	3866	GAACCCCTATTTGTTTATTTTCTTAAATACATTCABAATATGATCCGCTCATGAGACAT	3925
	/notes="initiation codon"	Db	2313	GAACCCCTATTTGTTTATTTTCTTAAATACATTCABAATATGATCCGCTCATGAGACAT	2254
	/codon start=1	Qy	3926	AACCTCGATAAATGCTTCAATAATATTGAAAAAGAGAGATGATGATTTCAACATTTCC	3985
	/transl_table=11	Db	2253	AACCTCGATAAATGCTTCAATAATATTGAAAAAGAGAGATGATGATTTCAACATTTCC	2194
	/protein_id="AAA66382.1"	Qy	3986	GTGTCGCCCTTATTCCTTTTTCGGCATTTTCGCCCTTCTGTTTTCCTCACCAGAAA	4045
	/db_xref="GI:806880"	Db	2193	GTGTCGCCCTTATTCCTTTTTCGGCATTTTCGCCCTTCTGTTTTCCTCACCAGAAA	2134
	/translations="MSIQHPRVALIPFAAFCLPVFAHPETLVKVKDAEDQLGARVGY	Qy	4046	CGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGACAGAGTGGGTATACATCGAAC	4105
	IELDLSKILESFRPEEPDMWTFKVLICGAVLSRI DAGQQLGR LHYSDNLVE	Db	2133	CGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGACAGAGTGGGTATACATCGAAC	2074
	DRPVEKHLTDGTVTRKELSAITMSDNTANLLLTITGGPKELTAFILHMGRVTRL	Qy	4106	TGGATCTCAACAGCGGTAAGATCTCTTGAGAGTTTTCGCCCGAAGAACGTTTTTCAATGA	4165
	LSALPAGNFIADKSGAGERSGRIIAALGPDGKPSRIVVITYTGSQATMDERNRQIA	Db	2073	TGGATCTCAACAGCGGTAAGATCTCTTGAGAGTTTTCGCCCGAAGAACGTTTTTCAATGA	2014
	EIGASLIKHW"	Qy	4166	TGAGCACTTTTAAAGTTCTGCTATGCGCGGCTTATTTATCCCGTATTTCAGCGCGGCAAG	4225
	/complement(2398..2853)	Db	2013	TGAGCACTTTTAAAGTTCTGCTATGCGCGGCTTATTTATCCCGTATTTCAGCGCGGCAAG	1954
	/standard_name="f1 origin"	Qy	4226	AGCAACTCGGTCCGCGCATACACTATTCTCAGAAATGACTTGGTTGAGTACTCACCAGTCA	4285
rep_origin	46.0%; Score 2677; DB 12; Length 3033;	Db	1953	AGCAACTCGGTCCGCGCATACACTATTCTCAGAAATGACTTGGTTGAGTACTCACCAGTCA	1894
ORIGIN	Best Local Similarity 100.0%; Pred. No. 0;				
	Matches 2677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	3146 TATAGTCAGTGTGTTATTACAAATTCACGCGCGTCTGTTTACAACTGCTGACTGGGAAAC				3205



```
QY 4286 CAGAAAGCATCTTACCGATGGCATGACAGTAAGAGAAATTATGCACTGCTGCCATACCA 4345
DB 1893 CAGAAAGCATCTTACCGATGGCATGACAGTAAGAGAAATTATGCACTGCTGCCATACCA 1834
QY 4346 TGAGTGATACACATCGCGCCCACTTACTCTCGAACAGATCGGAGGACCGAAGGCTAA 4405
DB 1833 TGAGTGATACACATCGCGCCCACTTACTCTCGAACAGATCGGAGGACCGAAGGCTAA 1774
QY 4406 CCGCTTTTTCGCAACAATGCGGGGATCATGTAACTCGCCCTTGATCTGTTGGGAACCGGAGC 4465
DB 1773 CCGCTTTTTCGCAACAATGCGGGGATCATGTAACTCGCCCTTGATCTGTTGGGAACCGGAGC 1714
QY 4466 TGAATGAAGCATACCAACGACGAGCGTGACACACCGATGCTCTGTAGCAATGGAACAA 4525
DB 1713 TGAATGAAGCATACCAACGACGAGCGTGACACACCGATGCTCTGTAGCAATGGAACAA 1654
QY 4526 CGTTGGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCGCGCAACAATTAATAG 4585
DB 1653 CGTTGGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCGCGCAACAATTAATAG 1594
QY 4586 ACTGGATGGAGGGGATAAAGTTGACGAGACCACTTCTGCGCTCGGCCCTTCCGCGCTGGCT 4645
DB 1593 ACTGGATGGAGGGGATAAAGTTGACGAGACCACTTCTGCGCTCGGCCCTTCCGCGCTGGCT 1534
QY 4646 GGTATTATGCTGAATAATCTGGAGCCGCTGAGCGGTGCTCGCGGTATCATTTGACGAC 4705
DB 1533 GGTATTATGCTGAATAATCTGGAGCCGCTGAGCGGTGCTCGCGGTATCATTTGACGAC 1474
QY 4706 TGGGGCCAGATGGTAAAGCCCTCCGATCTAGTTATCTACACGCGGGAGTCAGGCAA 4765
DB 1473 TGGGGCCAGATGGTAAAGCCCTCCGATCTAGTTATCTACACGCGGGAGTCAGGCAA 1414
QY 4766 CTATGATGAACGAAATAGACAGATCGCTGAGATAGTGGCTCACTGATTAAGCATTTGGT 4825
DB 1413 CTATGATGAACGAAATAGACAGATCGCTGAGATAGTGGCTCACTGATTAAGCATTTGGT 1354
QY 4826 AACTGTCAGAACCAAGTTTACTCATATATATCTTTAGATGATTAATACTTCAATTTTAAAT 4885
DB 1353 AACTGTCAGAACCAAGTTTACTCATATATATCTTTAGATGATTAATACTTCAATTTTAAAT 1294
QY 4886 TTAAGAGGATCTAGTCAAGATCTTTTTCATATCTCATGACCAAAATCCCTTAAAGTG 4945
DB 1293 TTAAGAGGATCTAGTCAAGATCTTTTTCATATCTCATGACCAAAATCCCTTAAAGTG 1234
QY 4946 AGTTTTCCTTCCACTGAGCGCTCAGACCCCGTAGAAAAGATCAAGAGATCTTCTTGAGATC 5005
DB 1233 AGTTTTCCTTCCACTGAGCGTCAAGCCCGTAGAAAAGATCAAGAGATCTTCTTGAGATC 1174
QY 5006 CTTTTTTTCTGCGGTAAATCTGCTGTGCAACAAACAAACCAACCGCTACGAGCGGTGG 5065
DB 1173 CTTTTTTTCTGCGGTAAATCTGCTGTGCAACAAACAAACCAACCGCTACGAGCGGTGG 1114
QY 5066 TTTGTTTTCGCGATCAAGAGCTACCAACTCTTTTTCGAAAGGTAACCTGCTTCAAGCAGAG 5125
DB 1113 TTTGTTTTCGCGATCAAGAGCTACCAACTCTTTTTCGAAAGGTAACCTGCTTCAAGCAGAG 1054
QY 5126 CGCAGATACCAAAATACCTGCTCTCTAGTGTAGCGGTAGTTAGGCCACCACTTCAAGAACT 5185
DB 1053 CGCAGATACCAAAATACCTGCTCTCTAGTGTAGCGGTAGTTAGGCCACCACTTCAAGAACT 994
QY 5186 CTGTAGACCGCGCTACATATCTGCTCTCTGCTATCTGTTTACCGTGGCTGCTGCCAGTG 5245
DB 993 CTGTAGACCGCGCTACATATCTGCTCTCTGCTATCTGTTTACCGTGGCTGCTGCCAGTG 934
QY 5246 GCGATTAAGTGTCTTCTTACCGGGTGGACTCAAGACGATAGTTACCGGATAAGCGCGAGC 5305
DB 933 GCGATTAAGTGTCTTCTTACCGGGTGGACTCAAGACGATAGTTACCGGATAAGCGCGAGC 874
QY 5306 GGTTCGGGCTGAAACGGGGGGTTCGTGTCACACAGCCGACCTTGGAGCGAAGCACTTACACCG 5365
DB 873 GGTTCGGGCTGAAACGGGGGGTTCGTGTCACACAGCCGACCTTGGAGCGAAGCACTTACACCG 814
```

```
QY 5366 AACTGAGATACCTACAGCGTGAGCTATGAGAAAGCCCGCTTCCGAGGGGAGAGG 5425
DB 813 AACTGAGATACCTACAGCGTGAGCTATGAGAAAGCCCGCTTCCGAGGGGAGAGG 754
QY 5426 CGGACAGGTATCCGGTAAAGCGGAGGGTCCGNAACAGGAGAGCGCACGAGGGAGCTTCCAG 5485
DB 753 CGGACAGGTATCCGGTAAAGCGGAGGGTCCGNAACAGGAGAGCGCACGAGGGAGCTTCCAG 694
QY 5486 GGGGAAACGCTCGGTATCTTTTATAGTCTCTGCGGGTTTCGCCACCTCTGACTTGAGCGTC 5545
DB 693 GGGGAAACGCTCGGTATCTTTTATAGTCTCTGCGGGTTTCGCCACCTCTGACTTGAGCGTC 634
QY 5546 GATTTTGTGATGCTCTGTCAGGGGGGCGGAGCTTATGGAAGAACGCCAGCAACCGGCGCT 5605
DB 633 GATTTTGTGATGCTCTGTCAGGGGGGCGGAGCTTATGGAAGAACGCCAGCAACCGGCGCT 574
QY 5606 TTTTACGGTTCTGCGCCCTTTTGTGCGCTTTTGTCTCACATGTTCTTTCCTGCTTATCCC 5665
DB 573 TTTTACGGTTCTGCGCCCTTTTGTGCGCTTTTGTCTCACATGTTCTTTCCTGCTTATCCC 514
QY 5666 CTGATCTGTGGATAACCGGTATTATACCGCTTTTGTAGTGAGCTGATACCGCTCGCGCAGCC 5725
DB 513 CTGATCTGTGGATAACCGGTATTATACCGCTTTTGTAGTGAGCTGATACCGCTCGCGCAGCC 454
QY 5726 GAACGACGAGCCGAGCGAGTCACTGAGCGAGGAGCGGAGAGCGGCCCATATCGCAAC 5785
DB 453 GAACGACGAGCCGAGCGAGTCACTGAGCGAGGAGCGGAGAGCGGCCCATATCGCAAC 394
QY 5786 CGCTCTCTCCCGCGCTTGGCGGATTCATTAATGACAG 5822
DB 393 CGCTCTCTCCCGCGCTTGGCGGATTCATTAATGACAG 357
```

## RESULT 15

```
AF041426/c 7823 bp DNA circular SYN 07-MAR-2000
LOCUS AF041426
DEFINITION Cloning vector pVLH-1, complete sequence.
ACCESSION AF041426
VERSION AF041426.1 GI:3025714
KEYWORDS Cloning vector pVLH-1
SOURCE Cloning vector pVLH-1
ORGANISM Cloning vector pVLH-1
REFERENCE 1 (bases 1 to 7823)
AUTHORS Deitsch,K.W. and Wellems,T.E.
TITLE Direct Submission
JOURNAL Submitted (08-JAN-1998) LPD/NIAID, NIH, Bldg 4, Rm BI-34, Bethesda, MD 20892, USA
FEATURES
    source
        1..7823
            /organism="Cloning vector pVLH-1"
            /mol_type="genomic DNA"
            /specific_host="Plasmodium falciparum"
            /db_xref="taxon:74888"
            /focus
        52..2651
            /organism="Plasmodium falciparum"
            /mol_type="genomic DNA"
            /strain="Pd2"
            /db_xref="taxon:5833"
            /chromosome="12"
        52..2651
            /genes="var7b"
        52..2651
            /genes="var7b"
            /note="variant surface antigen regulatory region"
        2653..4311
            /codon_start=1
            /note="expression marker"
        /product="luciferase"
        /protein_id="AAC12726.1"
        /db_xref="GI:3025715"
        /translation="MLHEDAKNIKKGPPPPPLEDGTAGBQLHKMKRYALVPQTIAF
```

[illegible]



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 05:09:12 ; Search time 1983 Seconds  
(without alignments)  
12472.521 Million cell updates/sec

Title: US-09-924-197-1  
Perfect score: 5822  
Sequence: 1 ctggcgcagcaggtttcccg.....tgccgattcattaatgcag 5822

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001as.\*
- 5: Geneseq2001bs.\*
- 6: Geneseq2002as.\*
- 7: Geneseq2003as.\*
- 8: Geneseq2003bs.\*
- 9: Geneseq2003cs.\*
- 10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5822	100.0	5822	6	AAS20853 Expressio
2	2757.2	47.4	5534	2	AAT43137 pUMIGIT a
3	2686	46.1	4229	4	Ash25896 Genetic i
4	2683.4	46.1	3448	7	Abq84197 Vector pg
5	2683.4	46.0	3448	7	Ad50589 pGEMT18H
6	2680.4	46.0	3448	7	Abt14439 HCV envel
7	2675.4	46.0	3018	2	Aax29905 Plasmid p
8	2675.4	46.0	5919	4	Aad09980 pHPG-GUS
9	2670.2	45.9	3357	7	Acc44700 Plasmid p
10	2670.2	45.9	3357	7	Abt16597 Artificia
11	2668.6	45.8	4514	8	Ada94775 Plasmid p
12	2658.4	45.7	9359	6	Abk11039 pVDH636 v
13	2658.4	45.7	9359	6	Abk10687 Transform
14	2641.4	45.4	3404	8	Ada41728 Plasmid p
15	2634.4	45.2	3968	4	Adad0981 pHP70-1M
16	2634.4	45.2	4626	4	Adad0988 pHP-ohx
17	2630.4	45.2	3485	6	Aai172813 Expressio
18	2553	43.9	4965	2	Aav18742 Complete
19	2553	43.9	5109	2	Aav13846 Complete
20	2553	43.9	6196	2	Aav13168 Complete
21	2553	43.9	6243	2	Aav13841 Complete
22	2553	43.9	6503	2	Aav13169 Complete
23	2553	43.9	6959	2	Aav13165 Complete

C 24	2553	43.9	7379	2	AAV13176	Aav13176 Complete
C 25	2553	43.9	8792	2	AAV18745	Aav18745 Complete
C 26	2552	43.8	10078	6	ABQ73047	Abq73047 Tomato an
C 27	2551.8	43.8	37808	2	AAX02780	Aax02780 Vector pm
C 28	2551.4	43.8	8618	2	AAV18741	Aav18741 Complete
C 29	2550	43.8	4883	6	AAAD28233	Aad28233 CMV expre
C 30	2550	43.8	7287	2	AAV02042	Aav02042 Plasmid p
C 31	2549.4	43.8	2962	9	ADD14879	Add14879 Phage dis
C 32	2549.4	43.8	3057	9	ADD14892	Add14892 Phage dis
C 33	2549.4	43.8	3093	9	ADD14876	Add14876 Phage dis
C 34	2549	43.8	2958	4	AAS13970	Aas13970 Plasmid p
C 35	2549	43.8	2958	4	AAS13969	Aas13969 Plasmid p
C 36	2549	43.8	2958	4	AAS13971	Aas13971 Plasmid p
C 37	2549	43.8	3351	6	ABT08167	Abt08167 Recombina
C 38	2549	43.8	4754	6	ABT08175	Abt08175 Recombina
C 39	2549	43.8	4773	6	ABT08176	Abt08176 Recombina
C 40	2549	43.8	4831	6	ABT08194	Abt08194 Recombina
C 41	2549	43.8	6345	2	AAX88054	Aax88054 Plasmid p
C 42	2549	43.8	21329	7	AAAD48562	Aad48562 ChimERIC
C 43	2545.8	43.7	6046	2	AAV13845	Aav13845 Complete
C 44	2545.8	43.7	6245	2	AAV13844	Aav13844 Complete
C 45	2545.8	43.7	6448	2	AAV13843	Aav13843 Complete

ALIGNMENTS

RESULT 1  
AAS20853  
ID AAS20853 standard; DNA, 5822 BP.  
XX  
AC AAS20853;  
XX  
DT 05-JUN-2002 (first entry)  
XX  
DE Expression cassette, pFP-IRN1 useful for gene silencing.

XX Inhibition of gene expression; expression cassette; inverted repeat; IR;  
XX NOS; nopaline synthase; figwort mosaic virus promoter; FMV; hsp70; PG;  
XX plant heat shock 70; tomato; polygalacturonase; disease resistance;  
XX flavour; nutritional characteristic; plant; gene silencing; pFP-IRN1;  
XX regulating gene expression; mutant; ds.

XX Lycopersicon esculentum.

OS Agrobacterium tumefaciens.

OS Figwort mosaic virus.

OS Synthetic.

OS Chimeric.

XX WO200214472-A2.  
XX 21-FBB-2002.

XX 14-AUG-2001; 2001WO-US025538.  
XX 15-AUG-2000; 2000US-0225508P.

XX 07-AUG-2001; 2001US-00924197.  
XX (DNAP ) DNA PLANT TECHNOLOGY CORP.

XX Guttersen N, Oeller P;

XX WPI; 2002-257599/30.

XX Reducing the expression of a target gene in a cell, comprises expressing

XX in the cell an expression cassette comprising a promoter operably linked

XX to a sense or antisense targeting sequence and an inverted repeat of a

XX nopaline synthase gene.

XX Claim 51; Page 37-39; 39pp; English.

XX The present invention relates to an improved method for inhibiting the

XX expression of a target gene in a cell, by expressing in the cell an

expression cassette comprising a promoter operably linked to a sense or antisense targeting sequence having substantial identity to a subsequence of the target gene, and an inverted repeat (IR) of a subsequence of a NOS (neopline synthase) gene, where the IR is unrelated to the targeting sequence. The expression cassette, pPP-IRN1 is constructed using a figwort mosaic virus (FMV) promoter in which the 5'-untranslated leader (5'-UTL) is derived from a plant heat shock 70 (hsp70) gene, the tomato polygalacturonase (PG) gene and an IR of the terminator of the Agrobacterium tumefaciens NOS gene. The method is useful for regulating expression of endogenous genes and transgenes, e.g. to regulate expression of endogenous plant phenotypes such as disease resistance, flavour, protein or nutritional characteristics. The improved gene silencing construct is used in functional genomics to determine the effect of regulating gene expression of a selected endogenous gene or transgene. The method is simple and rapid, and is suitable for high-throughput studies. Multiple transgenic constructs all containing the same repeat element can be silenced at the same time, since the initial silencing trigger mediated through the inverted repeat region will apply to all of the transcripts. The present sequence represents the expression cassette, pPP-IRN1

XX Sequence 5822 BP; 1689 A; 1275 C; 1292 G; 1566 T; 0 U; 0 Other;

Query Match 100.0%; Score 5822; DB 6; Length 5822;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGCAGCAGAGTTTCCGACCTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAG 60  
DB 1 CTGGCAGCAGAGTTTCCGACCTGGAAAGCGGCGAGTGAGCGCAATTAATGTGAG 60  
QY 61 TTAGCTCACTAATAGGCAACCCAGGCTTTACATTTTATGCTTCCGGCTGTAATGTG 120  
DB 61 TTAGCTCACTAATAGGCAACCCAGGCTTTACATTTTATGCTTCCGGCTGTAATGTG 120  
QY 121 TGAATTTGTAGCGGTAACAAATTTTACACAGGAACAGCTATGACCATGATTAACGCCAA 180  
DB 121 TGAATTTGTAGCGGTAACAAATTTTACACAGGAACAGCTATGACCATGATTAACGCCAA 180  
QY 181 GCTATTTAGTGACACTATAGTAATCTCAAGCTATGATCTCAAGCTTGGGAGCTCTCC 240  
DB 181 GCTATTTAGTGACACTATAGTAATCTCAAGCTATGATCTCAAGCTTGGGAGCTCTCC 240  
QY 241 CATATGGTCCAGCTCGAGCGGCGGCACTAGTGATCTTAGATCTCGAGTGGAGCTAAT 300  
DB 241 CATATGGTCCAGCTCGAGCGGCGGCACTAGTGATCTTAGATCTCGAGTGGAGCTAAT 300  
QY 301 TCTCAGTCCAAAGCCTCAACAGGTCAGGGTACAGATCTCCAAACATTTAGCCAAAGC 360  
DB 301 TCTCAGTCCAAAGCCTCAACAGGTCAGGGTACAGATCTCCAAACATTTAGCCAAAGC 360  
QY 361 TACAGGAGATCAATGAAGAACTTCAATCAAGTAAACTACTGTTCCAGCAGATGCATCA 420  
DB 361 TACAGGAGATCAATGAAGAACTTCAATCAAGTAAACTACTGTTCCAGCAGATGCATCA 420  
QY 421 TGGTCAAGTAAATTTGAGAAAGACATCCACCGAGACTTAAAGTTAGTGGGATCTTTG 480  
DB 421 TGGTCAAGTAAATTTGAGAAAGACATCCACCGAGACTTAAAGTTAGTGGGATCTTTG 480  
QY 481 AAAGTAATCTTGTCAACATCGAGCAGCTGGCTTGGGGACCGACAGCAAAAAGGAATGGT 540  
DB 481 AAAGTAATCTTGTCAACATCGAGCAGCTGGCTTGGGGACCGACAGCAAAAAGGAATGGT 540  
QY 541 GCAGAAATTTAGGCGCCTTACCAAAAGCATCTTTGCGCTTTTATTCGAAAGATAAAGCAG 600  
DB 541 GCAGAAATTTAGGCGCCTTACCAAAAGCATCTTTGCGCTTTTATTCGAAAGATAAAGCAG 600  
QY 601 ATTCTCTAGTACAAGTGGGGAACAAATTAACGTTGGAAAGAGCTGTCTGACAGCCAC 660  
DB 601 ATTCTCTAGTACAAGTGGGGAACAAATTAACGTTGGAAAGAGCTGTCTGACAGCCAC 660  
QY 661 TCCTAATGCTGATGACGAACGAGTGACGACCAACAAAGAAATTAAGTTCAGCTCAGGAT 720  
DB 661 TCCTAATGCTGATGACGAACGAGTGACGACCAACAAAGAAATTAAGTTCAGCTCAGGAT 720

DB 661 TCCTAATGCTGATGACGAACGAGTGACGACCAACAAAGAAATTAAGTTCAGCTCAGGAT 720  
QY 721 TTAGCAGCATTTCCAGATTGGGTTCAATCAACAAGGTACGAGCCATATACATTTATCAAA 780  
DB 721 TTAGCAGCATTTCCAGATTGGGTTCAATCAACAAGGTACGAGCCATATACATTTATCAAA 780  
QY 781 TTGGTATGCGCAAAACCAAGAGAACTCCCATCTCTCAAAAGGTTTGTAAAGGAATTTCT 840  
DB 781 TTGGTATGCGCAAAACCAAGAGAACTCCCATCTCTCAAAAGGTTTGTAAAGGAATTTCT 840  
QY 841 CAGTCCAAAGCCTCAACAGAGTACAGGTACAGAGTCTCCAAACCATTTAGCCAAAGCTAC 900  
DB 841 CAGTCCAAAGCCTCAACAGAGTACAGGTACAGAGTCTCCAAACCATTTAGCCAAAGCTAC 900  
QY 901 AGGAGATCAATGAAGAACTTTCAATCAAAAGTAACTACTGTTCCAGCAGCATCATGCG 960  
DB 901 AGGAGATCAATGAAGAACTTTCAATCAAAAGTAACTACTGTTCCAGCAGCATCATGCG 960  
QY 961 TCAGTAAAGTTTCAGAAAAGACATCCACCGAAGACTTAAAGTTAGTGGGCTCTTTGAAA 1020  
DB 961 TCAGTAAAGTTTCAGAAAAGACATCCACCGAAGACTTAAAGTTAGTGGGCTCTTTGAAA 1020  
QY 1021 GTAATCTGTCACATCGAGCAGCTGGCTTGTGGGACCGACAGCAAAAAGAAATGTTGCA 1080  
DB 1021 GTAATCTGTCACATCGAGCAGCTGGCTTGTGGGACCGACAGCAAAAAGAAATGTTGCA 1080  
QY 1081 GAATTTAGGCGCACCTACCAAAAGCATCTTTGCGCTTTATTTGCAAGATAAAGCAGATT 1140  
DB 1081 GAATTTAGGCGCACCTACCAAAAGCATCTTTGCGCTTTATTTGCAAGATAAAGCAGATT 1140  
QY 1141 CCTCTAGTACAGTGGGGAACAAATTAACGTTGAAAGAGCTGTCTCTGACAGCCACTCA 1200  
DB 1141 CCTCTAGTACAGTGGGGAACAAATTAACGTTGAAAGAGCTGTCTCTGACAGCCACTCA 1200  
QY 1201 CTAATCGGTATGACGAAACGAGTACGACCAACAAAGAAATTCCTCTATATAGAAAGCA 1260  
DB 1201 CTAATCGGTATGACGAAACGAGTACGACCAACAAAGAAATTCCTCTATATAGAAAGCA 1260  
QY 1261 TTCAATTTCCATTTTGAAGGACACAGAAAAATTTGCTACATTTTTCACAAATAT 1320  
DB 1261 TTCAATTTCCATTTTGAAGGACACAGAAAAATTTGCTACATTTTTCACAAATAT 1320  
QY 1321 TATTCAATTTTGTGACGCTTTCAAACTCTTTGTTTCTGTTTGTGATTTAGGAAATATT 1380  
DB 1321 TATTCAATTTTGTGACGCTTTCAAACTCTTTGTTTCTGTTTGTGATTTAGGAAATATT 1380  
QY 1381 AAAACCATGTTCTTAAAAACAGAAATTTATCTCTCAAGCAAAATCACCTTTTCAGGTCCA 1440  
DB 1381 AAAACCATGTTCTTAAAAACAGAAATTTATCTCTCAAGCAAAATCACCTTTTCAGGTCCA 1440  
QY 1441 TGCAGATCTTCTAATTTTCAAGTAAAGATTTTGGATCTTTAGAGCATCTAGTAAATTTCA 1500  
DB 1441 TGCAGATCTTCTAATTTTCAAGTAAAGATTTTGGATCTTTAGAGCATCTAGTAAATTTCA 1500  
QY 1501 GACTACAAAGATAGAGGCTTTGGATTTGCTTTGTATAGTTGTTCAAAATTTAGTTGTTGA 1560  
DB 1501 GACTACAAAGATAGAGGCTTTGGATTTGCTTTGTATAGTTGTTCAAAATTTAGTTGTTGA 1560  
QY 1561 GGAGGAGGAATCTAATGCGCAATGCAAGTATGTTGGCCCAAGTTCTTCCAAATTAAT 1620  
DB 1561 GGAGGAGGAATCTAATGCGCAATGCAAGTATGTTGGCCCAAGTTCTTCCAAATTAAT 1620  
QY 1621 AAATCACTGCCATGCAAGGATGCAACACCGGCTTAACTTTTGGAAATTTGCAAAATTTG 1680  
DB 1621 AAATCACTGCCATGCAAGGATGCAACACCGGCTTAACTTTTGGAAATTTGCAAAATTTG 1680  
QY 1681 AAAGTGAATTAATTAAGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740  
DB 1681 AAAGTGAATTAATTAAGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740  
QY 1741 ACTAATGTTGTAGCTTCAAAATTTGATGATCAATGCTTTCAGCAAGAGCCCAATCTGAT 1800  
DB 1741 ACTAATGTTGTAGCTTCAAAATTTGATGATCAATGCTTTCAGCAAGAGCCCAATCTGAT 1800



3961 Db ||||| AAGGATAGATATCAACATTTCCGTTGTCGCCCTTAATCCCTTTTGTGGCATTTTGC 4020  
4021 Qy ||||| CTTCTCTGTTTGTCTACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAAGTTG 4080  
4021 Db ||||| CTTCTCTGTTTGTCTACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAAGTTG 4080  
4081 Qy ||||| GGTGCACGAGTGGTTTACATCGAATCTGAACTGCTCAACAGCGTGAAGATCCTTGAGAGTTT 4140  
4081 Db ||||| GGTGCACGAGTGGTTTACATCGAATCTGAACTGCTCAACAGCGTGAAGATCCTTGAGAGTTT 4140  
4141 Qy ||||| CGCCCGGAGAACGTTTTCATATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTA 4200  
4141 Db ||||| CGCCCGGAGAACGTTTTCATATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTA 4200  
4201 Qy ||||| TTATCCCGGTTGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTTCTCAGAAT 4260  
4201 Db ||||| TTATCCCGGTTATGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTTCTCAGAAT 4260  
4261 Qy ||||| GACTTGGTTGAGTACTCACCAGTACACAGAAAGCATCTTACGGATGGCATGACAGTAAGA 4320  
4261 Db ||||| GACTTGGTTGAGTACTCACCAGTACACAGAAAGCATCTTACGGATGGCATGACAGTAAGA 4320  
4321 Qy ||||| GAATTATGCACTGCTGCCATAACCATGAGTGATAAACAACCTGCGCCTAACTTCTTGACA 4380  
4321 Db ||||| GAATTATGCACTGCTGCCATAACCATGAGTGATAAACAACCTGCGCCTAACTTCTTGACA 4380  
4381 Qy ||||| ACGATCGGAGGACCGAAGGAGCTAAACCGCTTTTGTGACAAACATGCGGAGTCAATGTAAC 4440  
4381 Db ||||| ACGATCGGAGGACCGAAGGAGCTAAACCGCTTTTGTGACAAACATGCGGAGTCAATGTAAC 4440  
4441 Qy ||||| CGCTTGTGATCGTGGGAAACCGAGCTGAATGAGCCATACCAACGACGAGCGGTGACACC 4500  
4441 Db ||||| CGCTTGTGATCGTGGGAAACCGAGCTGAATGAGCCATACCAACGACGAGCGGTGACACC 4500  
4501 Qy ||||| ACGATGCTCTAGCAATGCGCAACAACTGCGGCAAACTATTAACCTGCGCACTACTTACT 4560  
4501 Db ||||| ACGATGCTCTAGCAATGCGCAACAACTGCGGCAAACTATTAACCTGCGCACTACTTACT 4560  
4561 Qy ||||| CTAGCTTCCCGGCAACAACTAATAGACTGGATGGAGCGGATAAAGTTGCGAGGACCACTT 4620  
4561 Db ||||| CTAGCTTCCCGGCAACAACTAATAGACTGGATGGAGCGGATAAAGTTGCGAGGACCACTT 4620  
4621 Qy ||||| CTGCGCTCGGCGCTTCCGCTGCTGCTGTTTATGCTGATTAATCTGAGCGGTGAGCGT 4680  
4621 Db ||||| CTGCGCTCGGCGCTTCCGCTGCTGCTGTTTATGCTGATTAATCTGAGCGGTGAGCGT 4680  
4681 Qy ||||| GGGTCTCGCGGTATCATTTGACGCACTGGGCGCAGATGTTAAGCCCTCCCGTATCGTAGTT 4740  
4681 Db ||||| GGGTCTCGCGGTATCATTTGACGCACTGGGCGCAGATGTTAAGCCCTCCCGTATCGTAGTT 4740  
4741 Qy ||||| ATCTACACGACGGGAGTCAAGGCAACTATGATGAACGAATAAGACAGATCGCTGAGATA 4800  
4741 Db ||||| ATCTACACGACGGGAGTCAAGGCAACTATGATGAACGAATAAGACAGATCGCTGAGATA 4800  
4801 Qy ||||| GGTGCTCTACATGATTAAGCATTTGTAACTGTGACCAAGTCTTACTCATATATATCTTTAG 4860  
4801 Db ||||| GGTGCTCTACATGATTAAGCATTTGTAACTGTGACCAAGTCTTACTCATATATATCTTTAG 4860  
4861 Qy ||||| ATTGATTTAAACTTCAATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGAAT 4920  
4861 Db ||||| ATTGATTTAAACTTCAATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGAAT 4920  
4921 Qy ||||| CTATGACCAAAATCCCTTAACGAGTGTGTTTCCATCTGAGCGTCAAGCCCGGTAGAA 4980  
4921 Db ||||| CTATGACCAAAATCCCTTAACGAGTGTGTTTCCATCTGAGCGTCAAGCCCGGTAGAA 4980  
4981 Qy ||||| AAGATCAAAAGCATCTTCTGAGATCCCTTTTCTGCGCGTAAATCTGCTGTCGCAACA 5040  
4981 Db ||||| AAGATCAAAAGCATCTTCTGAGATCCCTTTTCTGCGCGTAAATCTGCTGTCGCAACA 5040  
5041 Qy ||||| AAAAAACACCGCTTACAGCGGTGTTTGTGTCGGGATCAAGAGCTACCAACTCTTTT 5100

5041 Db AAAAAACACCGCTTACAGCGGTGTTTGTGTCGGGATCAAGAGCTACCAACTCTTTT 5100  
5101 Qy CCGAAGGTAACTGGCTTTCAGCAGAGCGCAATACCAAAATATCTGTCTTCTAGTGTAGCGG 5160  
5101 Db CCGAAGGTAACTGGCTTTCAGCAGAGCGCAATACCAAAATATCTGTCTTCTAGTGTAGCGG 5160  
5161 Qy TAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCTCTGCTAATC 5220  
5161 Db TAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCTCTGCTAATC 5220  
5221 Qy CTGTATACAGTGGCTGCTGCCAGTGGCGAATAGTCTGTCTTACCGGGTTGGACTCAAGA 5280  
5221 Db CTGTATACAGTGGCTGCTGCCAGTGGCGAATAGTCTGTCTTACCGGGTTGGACTCAAGA 5280  
5281 Qy CGATAGTTACCGGATAGGCGCGCTCGGGCTGAAACGGGGGGTTCGTGCACACAGCCC 5340  
5281 Db CGATAGTTACCGGATAGGCGCGCTCGGGCTGAAACGGGGGGTTCGTGCACACAGCCC 5340  
5341 Qy AGCTTGGAGCGAAGCACTTACACCGAATCAGATACCTACAGCGTGAAGTATGAGAAGC 5400  
5341 Db AGCTTGGAGCGAAGCACTTACACCGAATCAGATACCTACAGCGTGAAGTATGAGAAGC 5400  
5401 Qy GCCACGCTTCCCAAGGCGAAGGCGGACAGGCTATCCGGTAAAGCGGCGAGGCTCGGAACA 5460  
5401 Db GCCACGCTTCCCAAGGCGAAGGCGGACAGGCTATCCGGTAAAGCGGCGAGGCTCGGAACA 5460  
5461 Qy GGAGAGCGCACAGAGGAGCTTCCAGGGGGGAAAACGCTTGTATCTTTATAGTCTGTCCGG 5520  
5461 Db GGAGAGCGCACAGAGGAGCTTCCAGGGGGGAAAACGCTTGTATCTTTATAGTCTGTCCGG 5520  
5521 Qy TTTGCGCACCTCTGACTTGAGCGTTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTA 5580  
5521 Db TTTGCGCACCTCTGACTTGAGCGTTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTA 5580  
5581 Qy TCGAAAAACCGCAGCAACCGCGCTTTTACGGTTCCTGCGCTTTTGTGCGCTTTTGTCT 5640  
5581 Db TCGAAAAACCGCAGCAACCGCGCTTTTACGGTTCCTGCGCTTTTGTGCGCTTTTGTCT 5640  
5641 Qy CACATGTTCTTCTCGTGTATCCCTGATCTCTGTGATTAACCGTATTTACCGCTTTTGA 5700  
5641 Db CACATGTTCTTCTCGTGTATCCCTGATCTCTGTGATTAACCGTATTTACCGCTTTTGA 5700  
5701 Qy TGAGCTGATACCGCTCGCGCAGCGCAACCGACGAGCGAGCTCAGTGCAGCGAGGAA 5760  
5701 Db TGAGCTGATACCGCTCGCGCAGCGCAACCGACGAGCGAGCTCAGTGCAGCGAGGAA 5760  
5761 Qy GCGAAGAGCGCCCAATACGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGC 5820  
5761 Db GCGAAGAGCGCCCAATACGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGC 5820  
5821 Qy AG 5822  
5821 Db AG 5822

RESULT 2  
AAT43137  
ID AAT43137 standard; DNA; 5534 BP.  
XX  
AC AAT43137;  
XX  
DT 10-FEB-1997 (first entry)  
XX  
DE pUMIGIT sequence including upstream activating sequence.  
XX  
KW Transgenic plant; gene expression; upstream activating sequence; UAS;  
KW transactivating protein; Gal4; herbicide resistance; polyhydroxybutyrate;  
XX safety; pUMIGIT; ds.  
OS Synthetic.  
XX  
PN CA2150039-A.  
XX



PD 09-AUG-1996.  
XX  
PF 24-MAY-1995; 95CA-02150039.  
XX  
PR 08-FEB-1995; 95GB-00002456.  
XX  
PA (UYWA-) UNIV WARWICK.  
XX  
PI Bennett M, May S, Ramsay N;  
XX  
DR WPI; 1996-486150/49.  
XX  
PT Control of genes in transgenic plants - using an upstream activating  
PT sequence activated by a transactivating protein expressed using a  
PT separate promoter.  
XX  
PS Example; Fig 5P; 48pp; English.  
XX  
CC Reporter plasmid pUMIGIT (AAV43137), or pUAS Minimal promoter Gus Int.  
CC Terminator, contains a beta-glucuronidase (GUS) reporter gene under the  
CC control of the 468 cauliflower-mosaic virus minimal promoter and 10  
CC synthetic 17-bp GAL4 binding sites. The upstream activating sequence  
CC (UAS) of pUMIGIT is activatable by yeast transactivating protein GAL4. In  
CC a novel method for controlling gene expression, a first transgenic plant  
CC carrying a gene encoding a desired phenotype (herbicide resistance of  
CC polyhydroxybutyrate prodn.) operatively linked to a UAS recognition site  
CC is pollinated by a second transgenic plant carrying Gal4 DNA (see also  
CC AAV43136). The transgene is fully expressed in F1 hybrid plants but  
CC segregates apart in subsequent generations, improving environmental  
CC safety  
XX  
SQ Sequence 5534 BP; 1400 A; 1355 C; 1387 G; 1363 T; 0 U; 29 Other;

Query Match  
Best Local Similarity 47.4%; Score 2757.2; DB 2; Length 5534;  
Matches 2891; Conservative 96.3%; Pred. No. 0;  
Mismatches 37; Indels 74; Gaps 4;

Qy 2840 GATCGTTCAACATTTGGCAATAAAGTTCTTAAGATTGAATCTTGTGGCCGCTTTCGG 2899  
Db 2366 GATCGTTCAACATTTGGCAATAAAGTTCTTAAGATTGAATCTTGTGGCCGCTTTCGG 2425

Qy 2900 ATGATTATCATATAATTTCTGTGTAATACGTTTAAAGCATGTAATAATTAACATGTAATGC 2959  
Db 2426 ATGATTATCATATAATTTCTGTGTAATACGTTTAAAGCATGTAATAATTAACATGTAATGC 2485

Qy 2960 ATGAGGTATTTATGAGATGGTGTGTTTATGATTAGATGCGCAATTAATACATTTAATAC 3019  
Db 2486 ATGAGGTATTTATGAGATGGTGTGTTTATGATTAGATGCGCAATTAATACATTTAATAC 2545

Qy 3020 GCGATAGAAACAAATAATAGCGCGCAAACTAGGATAAATATTCGCGCGCGGTGTCATCT 3079  
Db 2546 GCGATAGAAACAAATAATAGCGCGCAAACTAGGATAAATATTCGCGCGCGGTGTCATCT 2605

Qy 3080 ATGTTACTAGATCGACC-----TGCAAGCATGGGATTCGATACCGTTCGACCTCGAGCGGG 3124  
Db 2606 ATGTTACTAGATCGNNNGAATTCGATATCAAGCTTATCGATACCGTTCGACCTCGAGCGGG 2665

Qy 3125 GACGTGCGGCGCCAAATTCGCCCTTATAGTGTGATTTAC---AAATTCAGTCCGCGTCTT 3181  
Db 2666 GCGCGGTACCAATTCGCCCTTATAGTGTGATTTACGCGCGCTCACTGCGCGTCTT 2725

Qy 3182 TTACAACTGCTGACTCGGGAACCCCTGCGGTACCCAACTTAATTCGCTTTGCGAGCACAT 3241  
Db 2726 TTACAACTGCTGACTCGGGAACCCCTGCGGTACCCAACTTAATTCGCTTTGCGAGCACAT 2785

Qy 3242 CCCCTTTCCGCGCTGAGTAAAGAGGCGCCGCAACCGATCGCCCTTCCCAACAG 3301  
Db 2786 CCCCTTTCCGCGCTGAGTAAAGAGGCGCCGCAACCGATCGCCCTTCCCAACAG 2845

Qy 3302 TTGCGCAGCTGAATGGCGAAT-GGACGCGCCTGTAGCGCGCAATTAAGCGCGCGGT 3360  
Db 2846 TTGCGCAGCTGAATGGCGAATGGGACCGCGCCTGTAGCGCGCAATTAAGCGCGCGGT 2905

Qy 3361 GTGGTGGTTTACGGCAGCGGTGACCGCTACACTTGGCAGCGCCCTAGCGCGCTCCTTTC 3420  
Db 2906 GTGGTGGTTTACGGCAGCGGTGACCGCTACACTTGGCAGCGCCCTAGCGCGCTCCTTTC 2965

Qy 3421 GCTTTCTCCCTTCCCTTCTCGCCACAGTTCCGCGGCTTCCCGCTCAAGCTCTAAATCGG 3480  
Db 2966 GCTTTCTCCCTTCCCTTCTCGCCACAGTTCCGCGGCTTCCCGCTCAAGCTCTAAATCGG 3025

Qy 3481 GGGCTCCCTTTAGGGTTCCGATTTAGAGCTTTACGGCACCTCGACCGCAAAAACCTTGAT 3540  
Db 3026 GGGCTCCCTTTAGGGTTCCGATTTAGAGCTTTACGGCACCTCGACCGCAAAAACCTTGAT 3085

Qy 3541 TTGGGTGATGGTTCAAGTAGTGGGCATCGCCCTGATAGACGGTTTTCGCGCTTGAAG 3600  
Db 3086 TAGGGTGAATGGTTCAAGTAGTGGGCATCGCCCTGATAGACGGTTTTCGCGCTTGAAG 3145

Qy 3601 TTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACCTGGAAACAACACTCAACCT 3660  
Db 3146 TTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACCTGGAAACAACACTCAACCT 3205

Qy 3661 ATCTCGGTCTATTTCTTTGATTTATAAGGGATTTTTCGCGATTTTCGCGCTTATTTGTTAAAA 3720  
Db 3206 ATCTCGGTCTATTTCTTTGATTTATAAGGGATTTTTCGCGATTTTTCGCGCTTATTTGTTAAAA 3265

Qy 3721 AATGAGCTGATTTACAAATATTTAAACGGATTTTAAACAAATATTAACGTTTACAAT 3780  
Db 3266 AATGAGCTGATTTAAACAAATATTTAAACGGATTTTAAACAAATATTAACGTTTACAAT 3325

Qy 3781 TCGCCTGATGCGGTATTTCTCTCTTACGATCTGTGCGGTATTTTCACACCGCATACAGT 3840  
Db 3326 T-----AGGT 3330

Qy 3841 GGCACTTTTCGGGAAATGTGCGGAAACCCCTATTTGTTATTTTCTAAATACATTTCA 3900  
Db 3331 GGCACTTTTCGGGAAATGTGCGGAAACCCCTATTTGTTATTTTCTAAATACATTTCA 3390

Qy 3901 AATATGATTCGCTCATGAGACAATAACCTGATAAATGCTTCAATAATTTGAAAAAGG 3960  
Db 3391 AATATGATTCGCTCATGAGACAATAAACCCTGATAAATGCTTCAATAATTTGAAAAAGG 3450

Qy 3961 AAGAGTATGAGTATTTCAACATTTCCGTGCGCCCTTATTTCCCTTTTTCGCGCATTTTCG 4020  
Db 3451 AAGAGTATGAGTATTTCAACATTTCCGTGCGCCCTTATTTCCCTTTTTCGCGCATTTTCG 3510

Qy 4021 CTTCTGTTTTCGCTCAACCGAAGAAACGCTGTGTGAAGTAAAGATGCTGGAAGATCAGTTG 4080  
Db 3511 CTTCTGTTTTCGCTCAACCGAAGAAACGCTGTGTGAAGTAAAGATGCTGGAAGATCAGTTG 3570

Qy 4081 GGTCACAGAGTGGGTATCATCGAATCTGGATCTCAACAGGGTAAAGATCCTTCGAGAGTTT 4140  
Db 3571 GGTCACAGAGTGGGTATCATCGAATCTGGATCTCAACAGGGTAAAGATCCTTCGAGAGTTT 3630

Qy 4141 GCGCCGGAAGAACGTTTTCCAATGATGAGACCTTTTAAAGTTTCTGCTATGTGCGCGGTA 4200  
Db 3631 GCGCCGGAAGAACGTTTTCCAATGATGAGACCTTTTAAAGTTTCTGCTATGTGCGCGGTA 3690

Qy 4201 TTATCCCTTATTTGACCGCGGGCAAGAGCAACTCGGTGCGCGCATACATTTTCTCAGAT 4260  
Db 3691 TTATCCCTTATTTGACCGCGGGCAAGAGCAACTCGGTGCGCGCATACATTTTCTCAGAT 3750

Qy 4261 GACTTGGTTGAGTACTCAACAGTCAACAGAAAGCATCTTACGGATGGCATGACAGTAAGA 4320  
Db 3751 GACTTGGTTGAGTACTCAACAGTCAACAGAAAGCATCTTACGGATGGCATGACAGTAAGA 3810

Qy 4321 GAAATATCAGTGTGCTCAATACATGATGATTAACACTGCGCGCAACTTACTTCTGACA 4380  
Db 3811 GAAATATCAGTGTGCTCAATACATGATGATTAACACTGCGCGCAACTTACTTCTGACA 3870

Qy 4381 ACGATCGAGGACCGGAAGAGCTAACCGCTTTTTCGACAACATCGGGGATCATGTAACT 4440  
Db 3871 ACGATCGAGGACCGGAAGAGCTAACCGCTTTTTCGACAACATCGGGGATCATGTAACT 3930

Qy 4441 CGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACAGCAGCGGTGACACC 4500

Db 3931 |||||CGCTTGCATCGTTGGGAACCGGAGCTGATGAGCCATACCAACGACGAGCGGTGACACC||| 3990  
Qy 4501 ACAGTGCCTGTAGCAATGGCAACAACTGTGGCAAACTATTAACTGGCGAACTACTTACT 4560  
Db 3991 ACAGTGCCTGTAGCAATGGCAACAACTGTGGCAAACTATTAACTGGCGAACTACTTACT 4050  
Qy 4561 CTAGCTTCCGGCAACAACTAATAGACTGAGTGGAGGGGAGTAAGTTGACGAGCACTT 4620  
Db 4051 CTAGCTTCCGGCAACAACTAATAGACTGAGTGGAGGGGAGTAAGTTGACGAGCACTT 4110  
Qy 4621 CTGGCTTCGGCCCTTCGGCTCGCTGTTTATTCCTGATAACTCTGAGCGCGGTGAGCGT 4680  
Db 4111 CTGGCTTCGGCCCTTCGGCTCGCTGTTTATTCCTGATAACTCTGAGCGCGGTGAGCGT 4170  
Qy 4681 GGGTCTCGCGGTATCATTTGACAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTT 4740  
Db 4171 GGGTCTCGCGGTATCATTTGACAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTT 4230  
Qy 4741 ATCTACAGGACGGGAGTCAGGCAACTATGATGAACGAATAGACAGATCGCTGAGAT 4800  
Db 4231 ATCTACAGGACGGGAGTCAGGCAACTATGATGAACGAATAGACAGATCGCTGAGAT 4290  
Qy 4801 GGTCCCTCACTGATTAGCACTTGTAACTGTTCAGACCAAGTTTACTCATATATATCTTTAG 4860  
Db 4291 GGTCCCTCACTGATTAGCACTTGTAACTGTTCAGACCAAGTTTACTCATATATATCTTTAG 4350  
Qy 4861 ATTGATTTAAACCTCACTTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTTGATAAT 4920  
Db 4351 ATTGATTTAAACCTCACTTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTTGATAAT 4410  
Qy 4921 CTATGACCAAAATCCCTTAAGTAGTTTTCGTTTCCTCACTGAGCGTCAGACCCCGTAGAA 4980  
Db 4411 CTATGACCAAAATCCCTTAAGTAGTTTTCGTTTCCTCACTGAGCGTCAGACCCCGTAGAA 4470  
Qy 4981 AAGATCAAGAGATCTCTTGAGATCTCTTTTCTGGCGTAATCTGCTGTGCAACA 5040  
Db 4471 AAGATCAAGAGATCTCTTGAGATCTCTTTTCTGGCGTAATCTGCTGTGCAACA 4530  
Qy 5041 AAAAAACACCGCTACCAAGCGGTGGTTTGTTCGGGATCAAGAGTACCAACTCTTTTT 5100  
Db 4531 AAAAAACACCGCTACCAAGCGGTGGTTTGTTCGGGATCAAGAGTACCAACTCTTTTT 4590  
Qy 5101 CCGAAGTAACTGGCTTCAGCAGAGCGCAGATACCAATACTGTCTCTAGTGTAGCCG 5160  
Db 4591 CCGAAGTAACTGGCTTCAGCAGAGCGCAGATACCAATACTGTCTCTAGTGTAGCCG 4650  
Qy 5161 TAGTTAGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATC 5220  
Db 4651 TAGTTAGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATC 4710  
Qy 5221 CTGTTACAGTGGCTGCTCCAGTGGCGATAGTGTCTTTACCGGGTTGGACTCAAGA 5280  
Db 4711 CTGTTACAGTGGCTGCTCCAGTGGCGATAGTGTCTTTACCGGGTTGGACTCAAGA 4770  
Qy 5281 CGATAGTTACCGGATAGCGCCAGCGTCCGGCTGAAACGGGGGTTCGTGCACACAGCCC 5340  
Db 4771 CGATAGTTACCGGATAGCGCCAGCGTCCGGCTGAAACGGGGGTTCGTGCACACAGCCC 4830  
Qy 5341 AGCTTGGAGCAACGACCTACACCGAACTGAGATACCTACAGCGTGTAGTATGAGAAGC 5400  
Db 4831 AGCTTGGAGCAACGACCTACACCGAACTGAGATACCTACAGCGTGTAGTATGAGAAGC 4890  
Qy 5401 GCCACGCTTCCGAAGGAGAAAGCGGCAAGGTATCCGGTAAGCGGAGGGTCCGAACA 5460  
Db 4891 GCCACGCTTCCGAAGGAGAAAGCGGCAAGGTATCCGGTAAGCGGAGGGTCCGAACA 4950  
Qy 5461 GCAGAGCCACGAGGAGCTTCCAGGGGGAACCGCTGTATCTTTATAGTCTCTCTCGGG 5520  
Db 4951 GCAGAGCCACGAGGAGCTTCCAGGGGGAACCGCTGTATCTTTATAGTCTCTCTCGGG 5010  
Qy 5521 TTTCGCCACCTCTGACTTGGCGTCTGATTTTGTGATGCTGCTCAGGGGGCGGAGCTA 5580

Db 5011 TTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCTGTCAGGGGGCGGAGCCTA 5070  
Qy 5581 TGGAAAAACGCGACGAAACGCGGCCCTTTTACGGTTCCTTGGCCTTTTGTGCGCCTTTTGTCT 5640  
Db 5071 TGGAAAAACGCGACGAAACGCGGCCCTTTTACGGTTCCTTGGCCTTTTGTGCGCCTTTTGTCT 5130  
Qy 5641 CACATGTTCTTCTCTGCGTTATCCCTGATTTCTGTGGATAACCGTATTACCGCCTTTTGTGAG 5700  
Db 5131 CACATGTTCTTCTCTGCGTTATCCCTGATTTCTGTGGATAACCGTATTACCGCCTTTTGTGAG 5190  
Qy 5701 TGAGCTGATACCGCTCCGCGCAGCGAAACGACCGAGCGCAGCGAGTCAAGTGTGAGCGAGGAA 5760  
Db 5191 TGAGCTGATACCGCTCCGCGCAGCGAAACGACCGAGCGCAGCGAGTCAAGTGTGAGCGAGGAA 5250  
Qy 5761 GCGGAAGAGCGCCCAATACGCAAAACCGCCTCTCCCGCGCGCTTGGCGCGATTCATTAATGC 5820  
Db 5251 GCGGAAGAGCGCCCAATACGCAAAACCGCCTCTCCCGCGCGCTTGGCGCGATTCATTAATGC 5310  
Qy 5821 AG 5822  
Db 5311 AG 5312

RESULT 3  
AAH25896/C  
ID AAH25896 standard; DNA; 4229 BP.  
XX AAH25896;  
AC AC  
XX XX  
DT 24-AUG-2001 (first entry)  
XX XX  
DB Genetic information stability related oligonucleotide #5.  
XX XX  
KW Genetic information stability measurement; drug resistance; ds.  
XX XX  
OS Synthetic.  
XX XX  
PN JP2001087000-A.  
XX XX  
PD 03-APR-2001.  
XX XX  
PF 17-SEP-1999; 99JP-00264320.  
XX XX  
PR 17-SEP-1999; 99JP-00264320.  
XX XX  
PA (SENT-) SENTAN KAGAKU GIJUTSU INCUBATION CENT KK.  
XX XX  
PS WPI; 2001-360322/38.  
XX XX  
CC The present invention describes a method for the measurement of stability of genetic information in an animal cell in which a vector having a gene resistant to a first drug and a gene sensitive to a second drug is introduced into the animal cell. This can be used for the measurement of stability of genetic information in an animal cell. The present sequence is an oligonucleotide used in the exemplification of the invention  
XX XX  
SQ Sequence 4229 BP; 1108 A; 1004 C; 1036 G; 1081 T; 0 U; 0 Other;

Query Match 46.1%; Score 2686; DB 4; Length 4229;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2686; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3137 AATTGCGCCCTATAGTGTGATTCATTAATTCATCTGGCGCTGTTTACAAAGTCTGTGAC 3196  
Db 3157 AATTGCGCCCTATAGTGTGATTCATTAATTCATCTGGCGCTGTTTACAAAGTCTGTGAC 3098  
Qy 3197 TGGGAAAAACCTGCGGTTTACCACCACTTAATCGCCTTGCAGCACATCCCTTTTGGCCAGC 3256

Db 3097 TGGGAAACCTTGGCGTTACCAACTTAATCGCCTTGGCGACATCCCCCTTTGCCAGC 3038  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
3257 TGGCGTAATAGCGAAGAGCGCGCACCGATCGCCTTCCCAACAGTTGCGCAGCGCTGAAT 3316  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
3037 TGGCGTAATAGCGAAGAGCGCGCACCGATCGCCTTCCCAACAGTTGCGCAGCGCTGAAT 2978  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
3317 GGCGAATGGAAGCGCGCTGTAGCGCGCATTAAGCGCGCGCGGTGTGGTGTACGGCGCA 3376  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2977 GGCGAATGGAAGCGCGCTGTAGCGCGCATTAAGCGCGCGCGGTGTGGTGTACGGCGCA 2918  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
3377 CGGTGACCGCTACACTTCCGAGCGCCCTAGGCGCCGCTCCCTTTCGCTTCTTCCCTTCCCT 3436  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2917 GCGTGACCGCTACACTTCCGAGCGCCCTAGGCGCGCTCCCTTTCGCTTCTTCCCTTCCCT 2858  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
3437 TTCTCGCCACGTTGCGCGCTTTCGCCCTCAAGCTCTAAATGCGGGGCTCCCTTTPAGGGT 3496  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2857 TTCTCGCCACGTTGCGCGCTTTCGCCCTCAAGCTCTAAATGCGGGGCTCCCTTTPAGGGT 2798  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
3497 TCCGATTTAGAGCTTTACGGCACCTCGACCGCGCAAAAACCTTGATTTGGGTGATGGTTTAC 3556  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2797 TCCGATTTAGAGCTTTACGGCACCTCGACCGCGCAAAAACCTTGATTTGGGTGATGGTTTAC 2738  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
3557 GTAGTGGCCATCGCCCTGTAGAGCGTTTTCGCCCTTTCGAGGTTGAGTCCAGTTTCT 3616  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2737 GTAGTGGCCATCGCCCTGTAGAGCGTTTTCGCCCTTTCGAGGTTGAGTCCAGTTTCT 2678  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
3617 TTAATAGTGGACTCTTGTTCAAACTGGAACAACACTCAACCCCTATCTCGGTCTATTCTT 3676  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2677 TTAATAGTGGACTCTTGTTCAAACTGGAACAACACTCAACCCCTATCTCGGTCTATTCTT 2618  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
3677 TTGATTTAAGAGGATTTGCGGATTCGCGCTATTTGGTTTAAATAAGAGCTGATTTAAC 3736  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2617 TTGATTTAAGAGGATTTGCGGATTTGCGCTATTTGGTTTAAATAAGAGCTGATTTAAC 2558  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
3737 AAATAATTAACCGCAATTTTAAACAAATATTAACGTTTACAAATTCGCTCGATCGGTAT 3796  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2557 AAATAATTAACCGCAATTTTAAACAAATATTAACGTTTACAAATTCGCTCGATCGGTAT 2498  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
3797 TTTCTCTTACGCACTCTGTGGGTAATTTACACCGCATACAGGTGGGCACTTTTCGGGAA 3856  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2497 TTTCTCTTACGCACTCTGTGGGTAATTTACACCGCATACAGGTGGGCACTTTTCGGGAA 2438  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
3857 ATGTGCGGGAAACCCCTATTGTTTATTTTCTAAATACATTCAAATATGTATCGCTCA 3916  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2437 ATGTGCGGGAAACCCCTATTGTTTATTTTCTAAATACATTCAAATATGTATCGCTCA 2378  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
3917 TGAGACAATAACCCCTGATAAATGCTTCAATAATTTGAAAGGAAGATGAGTATTC 3976  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2377 TGAGACAATAACCCCTGATAAATGCTTCAATAATTTGAAAGGAAGATGAGTATTC 2318  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
3977 AACATTTCCGTGTGCGCCCTTATCCCTTTTTCGGGCAATTTGCTTCTCTGTTTTCCTC 4036  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2317 AACATTTCCGTGTGCGCCCTTATTCCTTTTTCGGGCAATTTGCTTCTCTGTTTTCCTC 2258  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4037 ACCCGAAGACCGTGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCGACAGTGGGTT 4096  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2257 ACCCGAAGACCGTGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCGACAGTGGGTT 2198  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4097 ACATCGAACTGATCTCAACAGCGTAAGATCTTTGAGAGTTTTCGCCCGCGAAGACGTT 4156  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2197 ACATCGAACTGATCTCAACAGCGTAAGATCTTTGAGAGTTTTCGCCCGCGAAGACGTT 2138  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4157 TTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTTATCCCGTATFGACG 4216  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2137 TTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTTATCCCGTATFGACG 2078  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4217 CCGGGCAAGAGAACTCGGTGCGCGCATACACTATTCTCAGAAAGCACTTGGTTGAGTACT 4276  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2077 CCGGGCAAGAGAACTCGGTGCGCGCATACACTATTCTCAGAAAGCACTTGGTTGAGTACT 2018  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4277 CACCAGTCACAGAAAGCACTTACGATGGCATGACAGTAAAGAAATTTATGACGTGCTG 4336  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 2017 CACCAGTCACAGAAAGCATCTTACGGATGGCATGACAGTAGAGAAATTTATGCAGTGCTG 1958  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4337 CCATAACCATGAGTGATTAACACTGCGCCAACTTACTTCTTGACAAACGATCGGAGGACGA 4396  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1957 CCATAACCATGAGTGATTAACACTGCGCCAACTTACTTCTTGACAAACGATCGGAGGACGA 1898  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4397 AGGAGCTAAACCGCTTTTTCACACATGCGGGGATCATGTAACCTCGCTTGTATCGTTGGG 4456  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1897 AGGAGCTAAACCGCTTTTTCACACATGCGGGGATCATGTAACCTCGCTTGTATCGTTGGG 1838  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4457 AACCGGAGCTGAATGAAGCCCATACCAACGACGAGCGTGACACACGATGCCCTGTAGCAA 4516  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1837 AACCGGAGCTGAATGAAGCCCATACCAACGACGAGCGTGACACACGATGCCCTGTAGCAA 1778  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4517 TGGCAACAAAGTTGCGCAAACTATTAACCTGCGCAACTTACTCTCTAGCTTCCCGGCAAC 4576  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1777 TGGCAACAAAGTTGCGCAAACTATTAACCTGCGCAACTTACTCTCTAGCTTCCCGGCAAC 1718  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4577 AATTAATAGACTGGATGGAGGGGATTAAGTTGACAGGACCACTTCTGCGCTCGGCCCTTC 4636  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1717 AATTAATAGACTGGATGGAGGGGATTAAGTTGACAGGACCACTTCTGCGCTCGGCCCTTC 1658  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4637 CGGCTGGCTGCTTATTTCTGATTAATCTGAGCCGCTGAGCGTGGGTCTCGCGGTATCA 4696  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1657 CGGCTGGCTGCTTATTTCTGATTAATCTGAGCCGCTGAGCGTGGGTCTCGCGGTATCA 1598  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4697 TTGAGCACTGCGGGCCAGATGTTAAGCTTCCGCTATCTGATGTTATCTACACGCGGGA 4756  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1597 TTGAGCACTGCGGGCCAGATGTTAAGCTTCCGCTATCTGATGTTATCTACACGCGGGA 1538  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4757 GTGAGCAACTATGAGTAAGCAAGCAATAGACAGATCGCTGAGATAGTGCCTCACTGATTA 4816  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1537 GTGAGCAACTATGAGTAAGCAAGCAATAGACAGATCGCTGAGATAGTGCCTCACTGATTA 1478  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4817 AGCATTTGCTAACTGTGACCAAGTTTACTCATATATATCTTTAGATTTGATTTAAACTTC 4876  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1477 AGCATTTGCTAACTGTGACCAAGTTTACTCATATATATCTTTAGATTTGATTTAAACTTC 1418  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4877 ATTTTAAATTTAAAGGATCTAGGTGAAGATCTCTTTTGAATATCTCATGACCAAAATCC 4936  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1417 ATTTTAAATTTAAAGGATCTAGGTGAAGATCTCTTTTGAATATCTCATGACCAAAATCC 1358  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4937 CTTAAGCTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAGGATCTT 4996  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1357 CTTAAGCTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAGGATCTT 1298  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4997 CTTGAGATCTCTTTTCTGCGGTAATCTGCTGCTTGCACAAACAAAACCAACCGCTAC 5056  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1297 CTTGAGATCTCTTTTCTGCGGTAATCTGCTGCTTGCACAAACAAAACCAACCGCTAC 1238  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
5057 CAGCGGTGTTTGTGTCGGATCAAGACTACCAACTCTTTTTCGAAAGGTAACTGGCT 5116  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1237 CAGCGGTGTTTGTGTCGGATCAAGACTACCAACTCTTTTTCGAAAGGTAACTGGCT 1178  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
5117 TCAGCAGACCGCAGATACCAATCTGCTTCTAGTGTAGCGGTAGTTAGGCCACCACT 5176  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1177 TCAGCAGACCGCAGATACCAATCTGCTTCTAGTGTAGCGGTAGTTAGGCCACCACT 1118  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
5177 TCAGAACTCTGTAGCAGCGCTCATATCTGCTCTCTGCTAACTCTTTTACAGTGGCTG 5236  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1117 TCAGAACTCTGTAGCAGCGCTCATATCTGCTCTCTGCTAACTCTTTTACAGTGGCTG 1058  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
5237 CTGCGAGTGGCGATAAGTCTGTCTTACCGGTTTGGACTCAAGAGATAGTTACCGGATA 5296  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1057 CTGCGAGTGGCGATAAGTCTGTCTTACCGGTTTGGACTCAAGAGATAGTTACCGGATA 998  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
5297 AGGCGCAGCGGTGCGGCTGAACCGGGGTTCTGTCACACAGCCAGCTTGGAGCGAACA 5356  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
997 AGGCGCAGCGGTGCGGCTGAACCGGGGTTCTGTCACACAGCCAGCTTGGAGCGAACA 938  
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
5357 CTTACACCGCAACTGAGATACCTTACAGCGTGTAGCTATGAGAAAGCGCCACGCTTCCGAG 5416  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
937 CTTACACCGCAACTGAGATACCTTACAGCGTGTAGCTATGAGAAAGCGCCAGCTTCCGAG 878

Qy 5417 GGAGAAAGCGGACAGGTATCCGGTAAGCGGACAGGTTCGGAACAGGAGAGCGACGAGGG 5476  
|||  
Db 877 GGAGAAAGCGGACAGGTATCCGGTAAGCGGACAGGTTCGGAACAGGAGAGCGACGAGGG 818  
|||  
Qy 5477 AGCTTCCAGGGGGAACGCGTGTATCTTTATAGTCTCTGTGCGGTTTGGCCACCTCTGAC 5536  
|||  
Db 817 AGCTTCCAGGGGGAACGCGTGTATCTTTATAGTCTCTGTGCGGTTTGGCCACCTCTGAC 758  
|||  
Qy 5537 TTGAGCGTCGATTTTGTGATGTCGTGTCAGGGGGCGGAGCCTATGGAAGAACGCCAGCA 5596  
|||  
Db 757 TTGAGCGTCGATTTTGTGATGTCGTGTCAGGGGGCGGAGCCTATGGAAGAACGCCAGCA 698  
|||  
Qy 5597 ACGGGGCTTTTACGGTTCTCGGCTTTCTGCGCTTTCTGCGCTTTGCTCAGATGTTCTTCTG 5656  
|||  
Db 697 ACGGGGCTTTTACGGTTCTCGGCTTTCTGCGCTTTCTGCGCTTTGCTCAGATGTTCTTCTG 638  
|||  
Qy 5657 CGTTATCCCTGATTCCTGTGATACCGTATTTACCGCTTTGAGTGAGCTGATACCGCTC 5716  
|||  
Db 637 CGTTATCCCTGATTCCTGTGATACCGTATTTACCGCTTTGAGTGAGCTGATACCGCTC 578  
|||  
Qy 5717 GCCGACCGGACACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAGCGGCCAA 5776  
|||  
Db 577 GCCGACCGGACACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAGCGGCCAA 518  
|||  
Qy 5777 TAGCAGAACCGCTCTCCCGCGGCTTGGCCGATTCATTATGCGAG 5822  
|||  
Db 517 TAGCAGAACCGCTCTCCCGCGGCTTGGCCGATTCATTATGCGAG 472  
|||

## RESULT 4

ABQ84197/c  
ID ABQ84197 standard; DNA; 3448 BP.

XX AC ABQ84197;

XX DT 19-FEB-2003 (first entry)

XX XX Vector pGEMTE1sh6 DNA SEQ ID NO:6.

XX DE Hepatitis C virus; HCV; envelope protein; hepatitis tropic; virucide;  
XX KW antiinflammatory; gene therapy; vaccine; gene; ds.

XX XX Hepatitis C virus.

XX OS Synthetic.

XX XX W0200285932-A2.

XX XX 31-OCT-2002.

XX PF 24-APR-2002; 2002MO-BE000062.

XX XX 24-APR-2001; 2001EP-00870088.

XX PR 17-JUL-2001; 2001US-0305604P.

XX XX (INNO-) INNOGENETICS NV.

XX XX Sablon E, Van Broekhoven A, Bosman A, Depla E, Deschamps G;

XX XX WPI; 2003-093095/08.

XX PT New recombinant nucleic acids for expressing Hepatitis C virus (HCV)  
XX PT envelope proteins in eukaryotic cells, comprising a sequence encoding a  
XX PT protein having an avian lysozyme leader peptide joined to the HCV  
XX PT envelope protein.

XX XX Example 1; Page 165-167; 319pp; English.

XX CC The present invention describes a recombinant nucleic acid (I) comprising  
XX CC a nucleotide sequence encoding a protein having an avian lysozyme leader  
XX CC peptide, or its functional equivalent, joined to a Hepatitis C virus  
XX CC (HCV) envelope protein or its part. Also described: (1) a vector  
XX CC comprising the recombinant nucleic acid; (2) a host cell comprising the

CC recombinant nucleic acid or the vector; and (3) a method for producing  
CC HCV envelope protein or its part in a host cell, comprising transforming  
CC the host cell with the recombinant nucleic acid or with the vector, where  
CC the host cell is capable of expressing the protein cited above. (I) has  
CC hepatotropic, virucide and antiinflammatory activities, and can be used  
CC in gene therapy and vaccines. The recombinant nucleic acid is useful for  
CC efficient expression of Hepatitis C virus envelope proteins in eukaryotic  
CC cells, such as yeast cells. The HCV envelope proteins may be used as a  
CC vaccine, or for incorporation into an immunoassay for the detection of  
CC anti-HCV antibodies, and/or genotyping of HCV, for prognosing or  
CC monitoring of HCV disease, or as a therapeutic agent. The method is used  
CC for producing HCV envelope protein or its part in a host cell. ABQ84197  
CC to ABQ84253 and ABP55528 to ABP55568 represent sequences used in the  
CC exemplification of the present invention  
XX

XX SQ Sequence 3448 BP; 862 A; 883 C; 879 G; 824 T; 0 U; 0 Other;

Query Match 46.1%; Score 2683.4; DB 7; Length 3448;

Best Local Similarity 99.4%; Pred. No. 0;  
Matches 2693; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 3114 CGGCCGATCGAGCTCGGGCCCAATTCGCCCTATAGTGTGCTATTCACATTCACCTGG 3173  
|||

Db 2985 CGCGAGCATCGAGCTCGGGCCCAATTCGCCCTATAGTGTGCTATTCACATTCACCTGG 2926  
|||

Qy 3174 CCGTGGTTTACACGCTGCTGCACTGGGAAACCTTGGGTTTACCACTTAATCGCCTTG 3233  
|||

Db 2925 CCGTGGTTTACACGCTGCTGCACTGGGAAACCTTGGGTTTACCACTTAATCGCCTTG 2866  
|||

Qy 3234 CAGCACATCCCTTTCCGAGCTGGCTTAATAGCGAGGCGCGACCATCGCCCTT 3293  
|||

Db 2865 CAGCACATCCCTTTCCGAGCTGGCTTAATAGCGAGGCGCGACCATCGCCCTT 2806  
|||

Qy 3294 CCCAACAGTTGCGCAGCTCGAATGGCGCCCTGTAGCGCGCATTAAGCGC 3353  
|||

Db 2805 CCCAACAGTTGCGCAGCTCGAATGGCGCCCTGTAGCGCGCATTAAGCGC 2746  
|||

Qy 3354 GCGGGTGTGGTTAGCGCGAGCTGACCGCTACACTTGCAGCGCCCTAGCGCCGC 3413  
|||

Db 2745 GCGGGTGTGGTTAGCGCGAGCTGACCGCTACACTTGCAGCGCCCTAGCGCCGC 2686  
|||

Qy 3414 TCCTTTCGCTTCTTCTTCCCTTCTCGCCAGCTTCGCGGCTTTCGCCGTCAAGCTCT 3473  
|||

Db 2685 TCCTTTCGCTTCTTCTTCCCTTCTCGCCAGCTTCGCGGCTTTCGCCGTCAAGCTCT 2626  
|||

Qy 3474 AAATCGGGGCTCCCTTTAGGGTTCCGATTTAGAGCTTTACGGCACTTCGACCGCAAAA 3533  
|||

Db 2625 AAATCGGGGCTCCCTTTAGGGTTCCGATTTAGAGCTTTACGGCACTTCGACCGCAAAA 2566  
|||

Qy 3534 ACTTGATTTGGGTGATGTTTCAGTGTGGGCCATCGCCCTGATAGAGGTTTTCGCC 3593  
|||

Db 2565 ACTTGATTTGGGTGATGTTTCAGTGTGGGCCATCGCCCTGATAGAGGTTTTCGCC 2506  
|||

Qy 3594 TTTGACGTTGAGTCCACAGTTCTTTAATAGTGAAGCTTCTTCCAAACTGGAAACAACACT 3653  
|||

Db 2505 TTTGACGTTGAGTCCACAGTTCTTTAATAGTGAAGCTTCTTCCAAACTGGAAACAACACT 2446  
|||

Qy 3654 CAACCTTATCTCGCTCTATTTCTTTTGAATTTAAGGGAATTTGCGGATTTTCGGCTATTG 3713  
|||

Db 2445 CAACCTTATCTCGCTCTATTTCTTTTGAATTTAAGGGAATTTGCGGATTTTCGGCTATTG 2386  
|||

Qy 3714 GTTAAAAAATGAGCTGATTTTAAACAAATTTTAAACGCAATTTTAAACAAATTTTAAACGTT 3773  
|||

Db 2385 GTTAAAAAATGAGCTGATTTTAAACAAATTTTAAACGCAATTTTAAACAAATTTTAAACGTT 2346  
|||

Qy 3774 TACAAATTTCCGCTGATGCGGTATTTCTCTCTTACGCACTGTGCGGTATTTTACACCGCA 3833  
|||

Db 2325 TACAAATTTCCGCTGATGCGGTATTTCTCTCTTACGCACTGTGCGGTATTTTACACCGCA 2266  
|||

Qy 3834 TACAGTGGCACTTTTCGGGAAATGTGCGGGAACCCCTATTATTTTCTTAAT 3893  
|||

Db 2265 TACAGTGGCACTTTTCGGGAAATGTGCGGGAACCCCTATTATTTTCTTAAT 2206  
|||

3894 ACATTCAATATGATCGCTCATGAGACAATAACCTGATATAATGCTTCAATATATTG 3953  
|||||  
2205 ACATTCAATATGATCGCTCATGAGACAATAACCTGATATAATGCTTCAATATATTG 2146  
|||||  
3954 AAAAGGAAGATGAGTATTCAACATTTCCGTGCGCCCTTATTCCTTTTTCGCGC 4013  
|||||  
2145 AAAAGGAAGATGAGTATTCAACATTTCCGTGCGCCCTTATTCCTTTTTCGCGC 2086  
|||||  
4014 ATTTTGCCTTCCTGTTTGTCTACCCAGAAACCGTGTGTAAGATGCTGAAGA 4073  
|||||  
2085 ATTTTGCCTTCCTGTTTGTCTACCCAGAAACCGTGTGTAAGATGCTGAAGA 2026  
|||||  
4074 TCAGTTGGTGCACAGTGGTGTACATCGAATGATCTCAACAGCGGTAGATCTTTGA 4133  
|||||  
2025 TCAGTTGGTGCACAGTGGTGTACATCGAATGATCTCAACAGCGGTAGATCTTTGA 1966  
|||||  
4134 CAGTTTTCGCCCCGGAAGACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGGG 4193  
|||||  
1965 GAGTTTTCGCCCCGGAAGACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGGG 1906  
|||||  
4194 CGCGTATTATCCGTTATGAGCGCGGCAAGAGCAACTCGGTGCGCGATACACTATTC 4253  
|||||  
1905 TACACTATTATCCGTTATGAGCGCGGCAAGAGCAACTCGGTGCGCGCGGCGGTATTC 1846  
|||||  
4254 TCAGATGACTTGGTTGAGTACTCACCAGTCAAGAAAGCATCTTACGATGGCGATGAC 4313  
|||||  
1845 TCAGATGACTTGGTTGAGTACTCACCAGTCAAGAAAGCATCTTACGATGGCGATGAC 1786  
|||||  
4314 AGTAAGAGAAATATGCACTGCTGCCATAACCATGATGATTAACACTGCGGCGCAACTTACT 4373  
|||||  
1785 AGTAAGAGAAATATGCACTGCTGCCATAACCATGATGATTAACACTGCGGCGCAACTTACT 1726  
|||||  
4374 TCTGACAAACGATCGGAGACCGAAGAGCTAACCGCTTTTGTGCAACAACATGGGGGATCA 4433  
|||||  
1725 TCTGACAAACGATCGGAGACCGAAGAGCTAACCGCTTTTGTGCAACAACATGGGGGATCA 1666  
|||||  
4434 TGTAACCTGCTGATGCTGCGGACCGGAGCTGAATGAAGCATACCAACGACGAGCG 4493  
|||||  
1665 TGTAACCTGCTGATGCTGCGGACCGGAGCTGAATGAAGCATACCAACGACGAGCG 1606  
|||||  
4494 TGACACCAACGATGCTGAGCAATGGCAACACGTTGCGCAACCTATTAACTGGCGAACT 4553  
|||||  
1605 TGACACCAACGATGCTGAGCAATGGCAACACGTTGCGCAACCTATTAACTGGCGAACT 1546  
|||||  
4554 ACTTACTAGCTTCCCGGCAACAATTAATAGACTGAGTGGAGCGGATAAAGTTGCAGG 4613  
|||||  
1545 ACTTACTAGCTTCCCGGCAACAATTAATAGACTGAGTGGAGCGGATAAAGTTGCAGG 1486  
|||||  
4614 ACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGTTTATTTGCTGATTAATCTGGAGCGG 4673  
|||||  
1485 ACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGTTTATTTGCTGATTAATCTGGAGCCGG 1426  
|||||  
4674 TGAGCGTGGTCTCGCGTATCATTTGACAGACTGGGGCCAGATGGTAAAGCCCTCCCGTAT 4733  
|||||  
1425 TGAGCGTGGTCTCGCGTATCATTTGACAGACTGGGGCCAGATGGTAAAGCCCTCCCGTAT 1366  
|||||  
4734 CGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGATGAACGAATAAGACAGATGCG 4793  
|||||  
1365 CGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGATGAACGAATAAGACAGATGCG 1306  
|||||  
4794 TGAGTAGTGGCTGCTCATGATTAAGCATTTGATGTTACTGTGACGCAAGTTTACTCATATAT 4853  
|||||  
1305 TGAGTAGTGGCTGCTCATGATTAAGCATTTGATGTTACTGTGACGCAAGTTTACTCATATAT 1246  
|||||  
4854 ACTTTAGATTGATTTAAACCTTCAATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTT 4913  
|||||  
1245 ACTTTAGATTGATTTAAACCTTCAATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTT 1186  
|||||  
4914 TGATAATCTCATGACCAAAATCCCTTAACTGAGGTTTTCCTGAGCGGTGACAGCC 4973  
|||||  
1185 TGATAATCTCATGACCAAAATCCCTTAACTGAGGTTTTCCTGAGCGGTGACAGCC 1126  
|||||  
4974 CGTAGAAGATCAAGGATCTTCTTGAGATCCTTTTCTGCGCGTAAATCTGCTGCTT 5033  
|||||

1125 CGTAGAAGATCAAGGATCTTCTTGAGATCCTTTTCTGCGGTAAATCTGCTGCTT 1066  
|||||  
5034 GCAAAACAAAAACACACCGCTTACAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAAC 5093  
|||||  
1065 GCAAAACAAAAACACACCGCTTACAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAAC 1006  
|||||  
5094 TCTTTTTCGGAAGGTAACTGGCTTACAGAGCGGAGATACCAATATCTGCTCTTCTACT 5153  
|||||  
1005 TCTTTTTCGGAAGGTAACTGGCTTACAGAGCGGAGATACCAATATCTGCTCTTCTACT 946  
|||||  
5154 GTAGCGGTAGTTAGSCCACCCTTCAAGAACTCTGTAGCACCGCTTACATATCTGCTCT 5213  
|||||  
945 GTAGCGGTAGTTAGSCCACCCTTCAAGAACTCTGTAGCACCGCTTACATATCTGCTCT 886  
|||||  
5214 GCTAATCTGTTTACAGTGGCTGCTCCAGTGGCGATTAAGTCTGTCTTACCGGGTGA 5273  
|||||  
885 GCTAATCTGTTTACAGTGGCTGCTCCAGTGGCGATTAAGTCTGTCTTACCGGGTGA 826  
|||||  
5274 CTCAGAGCATAGTTTACCGGATAGGCGGAGCGGCTGCGGCTGAAACGGGGGTTCTGTCAC 5333  
|||||  
825 CTCAGAGCATAGTTTACCGGATAGGCGGAGCGGCTGCGGCTGAAACGGGGGTTCTGTCAC 766  
|||||  
5334 ACAGCCAGCTTGGAGCGAACGACCTTACACCGAACTGAGATACCTACAGCGTGAGCTATG 5393  
|||||  
765 ACAGCCAGCTTGGAGCGAACGACCTTACACCGAACTGAGATACCTACAGCGTGAGCTATG 706  
|||||  
5394 AGAAGCGCCACGCTTCCGGAAGGAGGAGGAGCGGACAGGTATCCGGTAAAGCGGCAAGGT 5453  
|||||  
705 AGAAGCGCCACGCTTCCGGAAGGAGGAGGAGCGGACAGGTATCCGGTAAAGCGGCAAGGT 646  
|||||  
5454 CGGAACAGAGAGCGACAGAGGAGCTTCCAGGGGGGAAACGCTGATCTTTATAGTCC 5513  
|||||  
645 CGGAACAGAGAGCGACAGAGGAGCTTCCAGGGGGGAAACGCTGATCTTTATAGTCC 586  
|||||  
5514 TGTGGGTTTCGCCACCTCTGACTTGGAGCTCGATTTTGTGATGCTCGTCAGGGGGGCG 5573  
|||||  
585 TGTGGGTTTCGCCACCTCTGACTTGGAGCTCGATTTTGTGATGCTCGTCAGGGGGGCG 526  
|||||  
5574 GAGCTATGAAAAACGCCAGCAACCGGCTTTTTCAGGTTCTGCGCTTTTGTGCGC 5633  
|||||  
525 GAGCTATGAAAAACGCCAGCAACCGGCTTTTTCAGGTTCTGCGCTTTTGTGCGC 466  
|||||  
5634 TTTTGTCTCAGATGTTCTTCTCGGTTATCCGCTGATTTCTGAGTAAACGTTATACCGC 5693  
|||||  
465 TTTTGTCTCAGATGTTCTTCTCGGTTATCCGCTGATTTCTGAGTAAACGTTATACCGC 406  
|||||  
5694 CTTTGTAGTGTGATGATACCGCTTCCGCGCAGCGCAACGACCGAGCGCAGTCAAGTGAG 5753  
|||||  
405 CTTTGTAGTGTGATGATACCGCTTCCGCGCAGCGCAACGACCGAGCGGAGTCAAGTGAG 346  
|||||  
5754 CGAGAAACGGAAGAGCGCCCAATAACGCAACCGCTTCTCCCGCGGCTTGGCGGATCA 5813  
|||||  
345 CGAGAAACGGAAGAGCGCCCAATAACGCAACCGCTTCTCCCGCGGCTTGGCGGATCA 286  
|||||  
5814 TTAATGAG 5822  
|||||  
285 TTAATGAG 277  
|||||

## RESULT 5

AAD50589/C

ID AAD50589 standard; DNA; 3448 BP.

XX

AC AAD50589;

XX

DT 24-MAR-2003 (first entry)

XX

DE pGEMTE1sh6 vector DNA.

XX

KW Hepatitis C virus; HCV; envelope protein; virucide; immunostimulant;

KW immune response; T-cell; therapy; infection; pharmaceutical; vaccine;

KW gene; da.



4674 TGAGCGTGGGTCTCGCGGTATCATTCAGCACTGGGCGCAGATGGTAGCCCTCCGGTAT 4733  
Db TGGCGTGGGTCTCGCGGTATCATTCAGCACTGGGCGCAGATGGTAGCCCTCCGGTAT 1366  
4734 CGTAGTTATCTACACGACGGGGAGTCAGCACTATGATGAAACGAAATAGACAGATCGC 4793  
Db CGTAGTTATCTACACGACGGGGAGTCAGCACTATGATGAAACGAAATAGACAGATCGC 1306  
4794 TGAGATAGGTGCTCACTGATTAAGCAATGTGTAACTGTTCAGACCAAGTTTACTCATATAT 4853  
Db TGAGATAGGTGCTCACTGATTAAGCAATGTGTAACTGTTCAGACCAAGTTTACTCATATAT 1246  
4854 ACTTTAGATTGATTTAAACTCTCACTTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTT 4913  
Db ACTTTAGATTGATTTAAACTCTCACTTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTT 1186  
4914 TGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCC 4973  
Db TGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCC 1126  
4974 CGTAGAAAGATCAAGAGATCTTCTTGAGATCCTTTTCTGCGCGTAAATCTGCTGCTT 5033  
Db CGTAGAAAGATCAAGAGATCTTCTTGAGATCCTTTTCTGCGCGTAAATCTGCTGCTT 1066  
5034 GCACCAACAAAACCAACCGCTACACGGGTGCTTTGTTGCGGATCAAGAGCTACCAAC 5093  
Db GCACCAACAAAACCAACCGCTACACGGGTGCTTTGTTGCGGATCAAGAGCTACCAAC 1006  
5094 TCTTTTCCGAAGGTAACTGCTTTCAGCAGACGCGAGATACCAATCTCTTCTAGT 5153  
Db TCTTTTCCGAAGGTAACTGCTTTCAGCAGACGCGAGATACCAATCTCTTCTAGT 946  
5154 GTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAACCGCTACATACCTCGCTCT 5213  
Db GTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAACCGCTACATACCTCGCTCT 886  
5214 GCTAATCTGTACAGTGGTCTGCTGCGAGTGGCGATAGTCTGTTTACCGGGTGGGA 5273  
Db GCTAATCTGTACAGTGGTCTGCTGCGAGTGGCGATAGTCTGTTTACCGGGTGGGA 826  
5274 CTCAGACGATAGTTACCGGATAAGGCGACGCTGCGGCTGAACGGGGTTCGTGCAC 5333  
Db CTCAGACGATAGTTACCGGATAAGGCGACGCTGCGGCTGAACGGGGTTCGTGCAC 766  
5334 ACAGCCAGCTTGGAGCGAACGACTTACACCGAATGAGATACCTACAGCGTGAAGTATG 5393  
Db ACAGCCAGCTTGGAGCGAACGACTTACACCGAATGAGATACCTACAGCGTGAAGTATG 706  
5394 AGAAGCGCCACGCTTCCGAGGAGGAGAAAGCGGACAGGTATCCGTTAAGCGCAGGT 5453  
Db AGAAGCGCCACGCTTCCGAGGAGGAGAAAGCGGACAGGTATCCGTTAAGCGCAGGT 646  
5454 CGGAACAGGAGCGCACGAGGAGCTTCCAGGGGGAACCGCTGGTATCTTTATAGTCC 5513  
Db CGGAACAGGAGCGCACGAGGAGCTTCCAGGGGGAACCGCTGGTATCTTTATAGTCC 586  
5514 TGTGCGGTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCGTACAGGGGGCG 5573  
Db TGTGCGGTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCGTACAGGGGGCG 526  
5574 GAGCCTATGAAAACCGCAGCAACGGGCTTTTACGGTTCTGCGCTTTTCTGCGCC 5633  
Db GAGCCTATGAAAACCGCAGCAACGGGCTTTTACGGTTCTGCGCTTTTCTGCGCC 466  
5634 TTTTGTCTCAGATGTTCTTCTGCGTTATCCCTGATTCTGTGATTAACCGTATTACCGC 5693  
Db TTTTGTCTCAGATGTTCTTCTGCGTTATCCCTGATTCTGTGATTAACCGTATTACCGC 406  
5694 CTTTGTAGTGAAGTATACCGCTCCCGCAGCGCAACCGAGCGCAGCGAGTCAAGTGAG 5753  
Db CTTTGTAGTGAAGTATACCGCTCCCGCAGCGCAACCGAGCGCAGCGAGTCAAGTGAG 346  
5754 CGAGGAAGCGGAGAGCGCCCATATACGAACCGCCTCTCTCCCGCGGTGTTGGCGGATCA 5813

Db 345 CGAGAGCGGAAGAGCGCCCATATAGCAACCGCTCTTCCCGCGGTTGGCGGATCA 286  
Qy 5814 TTAATGCAG 5822  
Db 285 TTAATGCAG 277  
RESULT 6  
ABT14439/c  
ID ABT14439 standard; DNA; 3448 BP.  
XX  
AC ABT14439;  
XX  
DT 18-FEB-2003 (first entry)  
XX  
DB HCV envelope protein related vector DNA sequence SEQ ID No 6.  
XX  
KW Hepatotropic; Hepatitis C Virus; HCV virus; HCV envelope protein; HCV E1;  
KW HCV E2; vaccine; Hansenula polymorpha; viral envelope protein;  
KW glycosylated; ds.  
XX  
OS Unidentified.  
XX  
PN WO200286100-A2.  
XX  
PD 31-OCT-2002.  
XX  
PP 24-APR-2002; 2002WO-BB000063.  
XX  
PR 24-APR-2001; 2001EP-00870088.  
PR 17-JUL-2001; 2001US-0305604P.  
XX  
PA (INNO-) INNOGENETICS NV.  
XX  
PI Bosman A, Depla B, Deschamps G, Sablon E, Samson I;  
PI Van Broekhoven A, Haelewyn J;  
XX  
DR WPI; 2003-093130/08.  
XX  
PT New Hepatitis C Virus (HCV) virus-like particle formed of an HCV envelope  
PT protein or its part, useful for treating or preventing HCV infection.  
XX  
PS Example 7; Page 158-160; 296pp; English.  
XX  
CC The invention relates to a Hepatitis C Virus (HCV) virus-like particle  
CC formed of an HCV envelope protein or its part, where the cysteines are  
CC chemically and reversibly modified. The HCV E1 and/or HCV E2 proteins are  
CC useful for the manufacture of a vaccine or medicament against HCV  
CC infection or for preparing a diagnostic kit. The Hansenula polymorpha is  
CC useful for the expression of viral envelope proteins that become  
CC glycosylated upon expression in this yeast species. This polynucleotide  
CC sequence represents the DNA of a vector relating to the HCV envelope  
CC proteins of the invention  
XX  
SQ Sequence 3448 BP; 861 A; 882 C; 879 G; 823 T; 0 U; 3 Other;  
Query Match 46.0%; Score 2680.4; DB 7; Length 3448;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 2690; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
Qy 3114 CGGCGCATGCGAGCTGCGGCCCAATTGCGCCCTATAGTACTGATTTACATTCACCTGG 3173  
Db 2985 CGGAGCATGCGAGCTGCGGCCCAATTGCGCCCTATAGTACTGATTTACATTCACCTGG 2926  
Qy 3174 CGGTGCTTTTACAACTGCTGAGCTGGGAAACCTGGGTTTACCAACTTAATCGCCTTG 3233  
Db 2925 CGGTGCTTTTACAACTGCTGAGCTGGGAAACCTGGGTTTACCAACTTAATCGCCTTG 2866  
Qy 3234 CAGCATATCCCCCTTTCGCCAGCTGGCGTAAATAGCGAAGAGCGCGACCGATGCCCTT 3293  
Db 2865 CAGCATATCCCCCTTTCGCCAGCTGGCGTAAATAGCGAAGAGCGCGACCGATGCCCTT 2806



QY 3294 CCACAGATTGGCAGCCTGAATGGCGAATGGAACGCGCCCTGTAGCGGCCCATTAAGCGC 3353  
DB 2805 CCACAGATTGGCAGCCTGAATGGCGAATGGAACGCGCCCTGTAGCGGCCCATTAAGCGC 2746  
QY 3354 GCGGGGTGTGGTGTACGCGCGACGCTGACCGGTACACTTGCAGCGCCCTAGCGCCGC 3413  
DB 2745 GCGGGGTGTGGTGTAGCGCGACGCTGACCGGTACACTTGCAGCGCCCTAGCGCCGC 2686  
QY 3414 TCCTTTGGCTTTCCTTCCCTTCTTTCGCGCAAGCTTTCGCGCGCTTTCCTCGTCAAGCTCT 3473  
DB 2685 TCCTTTGGCTTTCCTTCCCTTCTTTCGCGCAAGCTTTCGCGCGCTTTCCTCGTCAAGCTCT 2626  
QY 3474 AATCCGGGGCTCCCTTTAGGGTTCGATTTAGAGCTTTTACGCACTTCACGCGCAAAAA 3533  
DB 2625 AATCCGGGGCTCCCTTTAGGGTTCGATTTAGAGCTTTTACGCACTTCACGCGCAAAAA 2566  
QY 3534 ACTTGATTTGGGTGATGTTTCAAGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCC 3593  
DB 2565 ACTTGATTTGGGTGATGTTTCAAGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCC 2506  
QY 3594 TTTGACGTTGGAGTCCACGTTCTTTTAACTAGTGACTCTTGTTCAAACTCGGAACAACCT 3653  
DB 2505 TTTGACGTTGGAGTCCACGTTCTTTTAACTAGTGACTCTTGTTCCAAACTCGGAACAACCT 2446  
QY 3654 CAACCCCTATCTCGGTCTATCTTTTGAATTTTAAAGGATTTTCCCGATTTTCGGCTATTG 3713  
DB 2445 CAACCCCTATCTCGGTCTATCTTTTGAATTTTAAAGGATTTTTCGGCTATTG 2386  
QY 3714 GTTAAAAAATGAGCTGATTTTAAACAATTTTAAACGGCAATTTTAAACAATTTTAAACGTT 3773  
DB 2385 GTTAAAAAATGAGCTGATTTTAAACAATTTTAAACGGCAATTTTAAACAATTTTAAACGTT 2326  
QY 3774 TACAATTTCCCTGATGCGGTATTTCTCTTACGATCTGTGGGTATTTTCAACGGCA 3833  
DB 2325 TACAATTTCCCTGATGCGGTATTTCTCTTACGATCTGTGGGTATTTTCAACGGCA 2266  
QY 3834 TACAGGTGGCACTTTTCGGGAAATGTGCGGAAACCCCTATTGTTTATTTTCTAAAT 3893  
DB 2265 TACAGGTGGCACTTTTCGGGAAATGTGCGGAAACCCCTATTGTTTATTTTCTAAAT 2206  
QY 3894 ACATTTCAAAATATGATCCGCTCATAGACAATAACCCCTGATAAATGCTTCAATAATTTG 3953  
DB 2205 ACATTTCAAAATATGATCCGCTCATAGACAATAACCCCTGATAAATGCTTCAATAATTTG 2146  
QY 3954 AAAAAAGAGATGATGAGTATTCAACATTTCCGTGTCGCCCTTATTTCCCTTTTTCGGGC 4013  
DB 2145 AAAAAAGAGATGATGAGTATTCAACATTTCCGTGTCGCCCTTATTTCCCTTTTTCGGGC 2086  
QY 4014 ATTTTCCCTTCTGTTTTCCTCACCCAGAAAACGCTGGTGAAGTAAAGATGCTGAAGA 4073  
DB 2085 ATTTTCCCTTCTGTTTTCCTCACCCAGAAAACGCTGGTGAAGTAAAGATGCTGAAGA 2026  
QY 4074 TCAGTTGGGTGACAGAGTGGTTTACATCGAATCGAATCTCAACAGCGGTAAAGATCTTTGA 4133  
DB 2025 TCAGTTGGGTGACAGAGTGGTTTACATCGAATCGAATCTCAACAGCGGTAAAGATCTTTGA 1966  
QY 4134 GAGTTTTCGCCCGGAAGAGTTTTCGAATGATGAGCACTTTTAAAGTTCTGCTATGTGG 4193  
DB 1965 GAGTTTTCGCCCGGAAGAGTTTTCGAATGATGAGCACTTTTAAAGTTCTGCTATGTGA 1906  
QY 4194 CGGGTATTATCCGTTATTGACGCGCGGAAGAGCAACTCGGTGCGCGCATACCTATTTC 4253  
DB 1905 TACACTATTATCCGTTATTGACGCGCGGAAGAGCAACTCGGTGCGCGCGCGGTATTTC 1846  
QY 4254 TCAGAAATGACTGGTTGAGTACTCACCGAGTACAGAAAGCACTTTTACGAGTGGCATGAC 4313  
DB 1845 TCAGAAATGACTGGTTGAGTACTCACCGAGTACAGAAAGCACTTTTACGAGTGGCATGAC 1786  
QY 4314 AGTAAGAGAAATTAAGAGTGGCTGCGATACCAATGATGATGAACAATCGCGGCAACTTACT 4373  
DB 1785 AGTAAGAGAAATTAAGAGTGGCTGCGATACCAATGATGATGAACAATCGCGGCAACTTACT 1726  
QY 4374 TCTGACAAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTCGACAAACATCGGGGATCA 4433

DB 1725 TCTGACAAACGATCGGAGGACCGAAGGAGCTTAACCGCTTTTTCGACAAACATGGGGATCA 1666  
QY 4434 TGTAACTCGCTTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCG 4493  
DB 1665 TGTAACTCGCTTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCG 1606  
QY 4494 TGAACACGATCGCTGTAGCAATGCGCAACAAAGTTGGCCAAACTATTAACTGGGGAAT 4553  
DB 1605 TGAACACGATCGCTGTAGCAATGCGCAACAAAGTTGGCCAAACTATTAACTGGGGAAT 1546  
QY 4554 ACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGTGATGAGCGGATGAAGTTGCAGG 4613  
DB 1545 ACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGTGATGAGCGGATGAAGTTGCAGG 1486  
QY 4614 ACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGCTTTTATTTGCTGATAAATCTGAGAGCGG 4673  
DB 1485 ACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGCTTTTATTTGCTGATAAATCTGAGAGCGG 1426  
QY 4674 TGAAGCTGGGTCTGCGGGTATCAATTCAGCACTCTGGGCCAGATGCTGAAGCCCTCCCGTAT 4733  
DB 1425 TGAAGCTGGGTCTGCGGGTATCAATTCAGCACTCTGGGCCAGATGCTGAAGCCCTCCCGTAT 1366  
QY 4734 CGTAGTTTATCTACACGACGGGAGTCAAGGCAACTATGATGAACGAAATAGACAGATGCG 4793  
DB 1365 CGTAGTTTATCTACACGACGGGAGTCAAGGCAACTATGATGAACGAAATAGACAGATGCG 1306  
QY 4794 TGAATAGGTGCTCTGATTAAGCAATGCTAACTGTACAGCAACAGTTTACTCATATAT 4853  
DB 1305 TGAATAGGTGCTCTGATTAAGCAATGCTAACTGTACAGCAACAGTTTACTCATATAT 1246  
QY 4854 ACTTTAGATGATTTTAAACTTCAATTTTAAAGATCTAGTGAAGATCTCTTTT 4913  
DB 1245 ACTTTAGATGATTTTAAACTTCAATTTTAAAGATCTAGTGAAGATCTCTTTT 1186  
QY 4914 TGAATACTCATGACCAAAATCCCTTAAACGTGAGTTTTCGTCCACTGAGCGTCAGACCC 4973  
DB 1185 TGAATACTCATGACCAAAATCCCTTAAACGTGAGTTTTCGTCCACTGAGCGTCAGACCC 1126  
QY 4974 CGTAGAAAAAGATCAAAAGGATCTTTTGAGATCTCTTTTTCGCGGCTAATCTGCTGCTT 5033  
DB 1125 CGNNAAAAAGATCAAAAGGATCTTTTGAGATCTCTTTTTCGCGGCTAATCTGCTGCTT 1066  
QY 5034 GCAACCAAAAAACACACGCTACAGCGGTGTTTGTTCGCGGATCAAGAGCTTACCAC 5093  
DB 1065 GCAACCAAAAAACACACGCTACAGCGGTGTTTGTTCGCGGATCAAGAGCTTACCAC 1006  
QY 5094 TCTTTTTCGAGAGTAACTGGCTTCAGCAGAGCGCAGATACCAATACTGTCTCTTAGT 5153  
DB 1005 TCTTTTTCGAGAGTAACTGGCTTCAGCAGAGCGCAGATACCAATACTGTCTCTTAGT 946  
QY 5154 GTAGCGGTAGTTAGGCCACACCTTCAAGAACTCTGTAGCAACCGCTTACATACCTCGCTCT 5213  
DB 945 GTAGCGGTAGTTAGGCCACACCTTCAAGAACTCTGTAGCAACCGCTTACATACCTCGCTCT 886  
QY 5214 GCTTAATCTGTTTACAGTGTGCTGCTCCAGTGGCGATTAAGTCTGTTTACCGGGTTTGA 5273  
DB 885 GCTTAATCTGTTTACAGTGTGCTGCTCCAGTGGCGATTAAGTCTGTTTACCGGGTTTGA 826  
QY 5274 CTCAGACGATAGTTTACCGGATTAAGCGCGCAGCGCTGCGGCTGAACCGGGGTTTCTGTCAC 5333  
DB 825 CTCAGACGATAGTTTACCGGATTAAGCGCGCAGCGCTGCGGCTGAACCGGGGTTTCTGTCAC 766  
QY 5334 ACAGCCAGCTTGGAGGGAACGACCTTACACGAACTGTAGATATCTACAGCGTGAAGTATG 5393  
DB 765 ACAGCCAGCTTGGAGGGAACGACCTTACACGAACTGTAGATATCTACAGCGTGAAGTATG 706  
QY 5394 AGAAGCGCCACGCTTCCGAGGAGAAAGCGGAGAGTATCCGTTAAGCGGAGGT 5453  
DB 705 AGAAGCGCCACGCTTCCGAGGAGAAAGCGGAGAGTATCCGTTAAGCGGAGGT 646  
QY 5454 CGGAACAGGAGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGCTTTTATAGTCC 5513

Db 645 CGGACAGGAGCGCAGAGGAGCTTCAGGGGAAACGCCCTGCTATCTTTATAGTCC 586  
Qy 5514 TGTGGGTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTGTCAGGGGGCG 5573  
Db 585 TGTGGGTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTGTCAGGGGGCG 526  
Qy 5574 GAGCCTATGAAACAGCCAGCAAGCGGCTTTTACGGTTCTGCGCTTTTGTGCGCC 5633  
Db 525 GAGCCTATGAAACAGCCAGCAAGCGGCTTTTACGGTTCTGCGCTTTTGTGCGCC 466  
Qy 5634 TTTTGTCTCAGATGTTCTTCTGCTGATATCCCTGATTTCTGTGATTAACCGTATTACCGC 5693  
Db 465 TTTTGTCTCAGATGTTCTTCTGCTGATATCCCTGATTTCTGTGATTAACCGTATTACCGC 406  
Qy 5694 CTTTGTAGTGTGATACCGCTCGCGCAGCGAAGCGAGCGAGCGAGTCACTGAG 5753  
Db 405 CTTTGTAGTGTGATACCGCTCGCGCAGCGAAGCGAAGCGAGCGAGTCACTGAG 346  
Qy 5754 CGAGGAGCGAGAGCGCCCAATACCAACCGCTCTCCCGCGGTTGGCGGATTCA 5813  
Db 345 CGAGGAGCGAGAGCGCCCAATACCAACCGCTCTCCCGCGGTTGGCGGATTCA 286  
Qy 5814 TTAATGCAG 5822  
Db 285 TTAATGCAG 277

## RESULT 7

AAX29905/c

ID AAX29905 standard; DNA; 3018 BP.

XX AC AAX29905;

XX XX

DT 06-JUL-1999 (first entry)

XX DE Plasmid pGEM (RTM)-T Easy Vector.

XX XX

XX Oligonucleotide; hybridisation; amidated polypeptide; hormone; primer;  
KW amplification; rat; brain; cholecystokinin; CCK; amidation; hormone;  
KW neurohormone; physiological role; plasmid; ss.

XX OS Synthetic.

XX PN W09910361-A1.

XX XX

PD 04-MAR-1999.

XX PF 07-AUG-1998; 98WO-FR001767.

XX XX

XX 26-AUG-1997; 97FR-00010643.

XX PA (SCRC ) SCRAS SOC CONSEILS RECH &amp; APPL SCI.

XX PI Martinez J, Goze C;

XX DR WPI; 1999-204645/17.

XX XX

XX Oligonucleotides that hybridize to mRNA encoding precursors of amidated

XX PT hormones - useful for isolating new hormones.

XX PS Disclosure; Page 19-21; 27pp; French.

XX XX

XX The invention relates to single-stranded oligonucleotides of 9-42  
CC nucleotides that hybridize, under stringent or non-stringent conditions,  
CC with mRNA encoding precursors of amidated polypeptide hormones. This  
CC sequence represents the plasmid pGEM (RTM)-T Easy vector used to subclone  
CC fragments of amidated polypeptide hormone coding sequences. The  
CC oligonucleotides are used to identify new amidated polypeptide hormones  
CC by hybridisation to cDNA encoding their precursors. cDNA encoding the  
CC hormones, particularly neurohormones, can be expressed in microorganisms  
CC and used therapeutically or to study their physiological roles

XX Sequence 3018 BP; 774 A; 747 C; 767 G; 730 T; 0 U; 0 Other;

SQ

Query Match 46.0%; Score 2675.4; DB 2; Length 3018;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2676; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3146 TATAGTCAGTCGTATTAACAATTCTACCTGCGCGCTGTTTACCAACGTCGTGACCTGGGAAAC 3205  
Db 3018 TATAGTCAGTCGTATTAACAATTCTACCTGCGCGCTGTTTACCAACGTCGTGACCTGGGAAAC 2959  
Qy 3206 CTTGGCGTTTACCAACCTTAATGCGCTTGCAGCAGATCCCTTTCCGCCAGCTGGCGTAA 3265  
Db 2958 CTTGGCGTTTACCAACCTTAATGCGCTTGCAGCAGATCCCTTTCCGCCAGCTGGCGTAA 2899  
Qy 3266 AGCGAAGAGGCGCGCAGATCGCGCTTCCCAACGTTGCGCAGCTGAAATGGCGAATGG 3325  
Db 2898 AGCGAAGAGGCGCGCAGATCGCGCTTCCCAACGTTGCGCAGCTGAAATGGCGAATGG 2839  
Qy 3326 AGCGCGCTCTGAGCGCGCATTAAGCGCGGCTGCTGTTTACCGCGAGCTGACGCG 3385  
Db 2838 AGCGCGCTCTGAGCGCGCATTAAGCGCGGCTGCTGTTTACCGCGAGCTGACGCG 2779  
Qy 3386 CTACACTTGGCAGCGCTTACGCGCGCTCTTTCGCTTTCTTCCCTTCTTCTTCTCGCA 3445  
Db 2778 CTACACTTGGCAGCGCTTACGCGCGCTCTTTCGCTTTCTTCCCTTCTTCTTCTCGCA 2719  
Qy 3446 CGTTGCGCGCTTTCCTCGCTCAAGCTCTAATTCGCGCGCTCTTTCGCGCTTTCGCA 3505  
Db 2718 CGTTGCGCGCTTTCCTCGCTCAAGCTCTAATTCGCGCGCTCTTTCGCGCTTTCGCA 2659  
Qy 3506 GAGCTTTTACGCGCAGCTCGACCGCAAAACCTTGATTTGGGTGATGTTTCACTAGTGGC 3565  
Db 2658 GAGCTTTTACGCGCAGCTCGACCGCAAAACCTTGATTTGGGTGATGTTTCACTAGTGGC 2599  
Qy 3566 CATCGCTCTGATAGACGCTTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCG 3625  
Db 2598 CATCGCTCTGATAGACGCTTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCG 2539  
Qy 3626 GACTTTTGTTCGCGCAACCTGCAACCTATCTCGCTCTATTTCTTCTTCTTCTTCTTCT 3685  
Db 2538 GACTTTTGTTCGCGCAACCTGCAACCTATCTCGCTCTATTTCTTCTTCTTCTTCTTCT 2479  
Qy 3686 AAGGGATTTTGGCGATTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCT 3745  
Db 2478 AAGGGATTTTGGCGATTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCT 2419  
Qy 3746 ACAGCAATTTTACCAAAATATTAACGTTTACAAATTCGCGCTGATGCGGTATTTCTCT 3805  
Db 2418 ACAGCAATTTTACCAAAATATTAACGTTTACAAATTCGCGCTGATGCGGTATTTCTCT 2359  
Qy 3806 ACAGCTCTGCGGTATTTTCACCGCATACAGGTGGCATTTTTCGCGGGAATTTGCGCG 3865  
Db 2358 ACAGCTCTGCGGTATTTTCACCGCATACAGGTGGCATTTTTCGCGGGAATTTGCGCG 2299  
Qy 3866 GAACCCCTATTTTGTATTTTCTTAAATATTAATTAATTAATTAATTAATTAATTAAT 3925  
Db 2298 GAACCCCTATTTTGTATTTTCTTAAATATTAATTAATTAATTAATTAATTAATTAAT 2239  
Qy 3926 AACCTCTGATTAATGCTTCAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 3985  
Db 2238 AACCTCTGATTAATGCTTCAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 2179  
Qy 3986 GTGTGCGCTTATTTCTCTTTTTCGCGCATTTTTCGCTTCTTCTTCTTCTTCTTCTTCT 4045  
Db 2178 GTGTGCGCTTATTTCTCTTTTTCGCGCATTTTTCGCTTCTTCTTCTTCTTCTTCTTCT 2119  
Qy 4046 CGCTGCTGAAGTAAAGATGCTGAGATCAGTTTGGGTGCAACGAGTGGTTTACATCGAAC 4105  
Db 2118 CGCTGCTGAAGTAAAGATGCTGAGATCAGTTTGGGTGCAACGAGTGGTTTACATCGAAC 2059  
Qy 4106 TGGATCTCAACAGCGGTAAAGATCTTTCAGAGTTTTCGCGCGGAGAGAGCTTTTCTCAATGA 4165  
Db 2058 TGGATCTCAACAGCGGTAAAGATCTTTCAGAGTTTTCGCGCGGAGAGAGCTTTTCTCAATGA 1999

QY 4166 TGAGCACTTTTAAAGTTCTGCTATGTGGCCGGGTATATATCCCGTATTTAGCCGGGCAAG 4225  
DB 1998 TGAGCACTTTTAAAGTTCTGCTATGTGGCCGGGTATATATCCCGTATTTAGCCGGGCAAG 1939  
QY 4226 AGCAACTCGGTCCCGCATACACTATTCTCAGAAATGACTTGGTTAGTACTCACCAAGTCA 4285  
DB 1938 AGCAACTCGGTCCCGCATACACTATTCTCAGAAATGACTTGGTTAGTACTCACCAAGTCA 1879  
QY 4286 CAGAAAGCACTTTACCGATGGCATGACAGTAAGAGAAATTTATGCAAGTGTGCCATAACCA 4345  
DB 1878 CAGAAAGCACTTTACCGATGGCATGACAGTAAGAGAAATTTATGCAAGTGTGCCATAACCA 1819  
QY 4346 TGAGTGATTAACACTGCGGCCCAACTTACTTCTGACACAGATCGAGAGACCGAAGAGCTAA 4405  
DB 1818 TGAGTGATTAACACTGCGGCCCAACTTACTTCTGACACAGATCGAGAGACCGAAGAGCTAA 1759  
QY 4406 CCGCTTTTGTGCAACAATCGGGGATCATGTAACTGCGCTTGTGATGCTGGTGGAAACCGGAGC 4465  
DB 1758 CCGCTTTTGTGCAACAATCGGGGATCATGTAACTGCGCTTGTGATGCTGGTGGAAACCGGAGC 1599  
QY 4466 TGAATGAAGCCATACCAACGACGAGCGTGACACCAAGATGCGCTGTAGCAATGCAACAA 4525  
DB 1698 TGAATGAAGCCATACCAACGACGAGCGTGACACCAAGATGCGCTGTAGCAATGCAACAA 1639  
QY 4526 CGTTGGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAG 4585  
DB 1638 CGTTGGCGAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAG 1579  
QY 4586 ACTGATGAGGCGGATTAAGTTGCGAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCT 4645  
DB 1578 ACTGATGAGGCGGATTAAGTTGCGAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCT 1519  
QY 4646 GGTATTATGCTGATAAATCTGAGCGCGTGAGCGTGCGCTCTGCGGTATCAATTCGACAC 4705  
DB 1518 GGTATTATGCTGATAAATCTGAGCGCGTGAGCGTGCGCTCTGCGGTATCAATTCGACAC 1459  
QY 4706 TGGGGCCAGATGTTAGCCCTCCGTTATCGTAGTTATCTACAGACGGGAGTCAGGCAA 4765  
DB 1458 TGGGGCCAGATGTTAGCCCTCCGTTATCGTAGTTATCTACAGACGGGAGTCAGGCAA 1399  
QY 4766 CTATGATGAACGAATPAGACAGATCGCTGAGTAGTGTCTCAGTCAATTAAGCATTTGT 4825  
DB 1398 CTATGATGAACGAATPAGACAGATCGCTGAGTAGTGTCTCAGTCAATTAAGCATTTGT 1339  
QY 4826 AACTGTGACACCAAGTTTACTCATATATCTTTAGATTGATTTAAATCTTCAATTTTAAAT 4885  
DB 1338 AACTGTGACACCAAGTTTACTCATATATCTTTAGATTGATTTAAATCTTCAATTTTAAAT 1279  
QY 4886 TTAAGGATCTAGGTGAAGATCGTTTGTGATATCTCATGACCAAAATCCCTTAAGCTG 4945  
DB 1278 TTAAGGATCTAGGTGAAGATCGTTTGTGATATCTCATGACCAAAATCCCTTAAGCTG 1219  
QY 4946 AGTTTGTCTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAAGATCTTCTTGAGATC 5005  
DB 1218 AGTTTGTCTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAAGATCTTCTTGAGATC 1159  
QY 5006 CTTTTTTTCTGCGCGTAATCTGCTGTGCAACAAAAACCAACCGCTACCGCGGTGG 5065  
DB 1158 CTTTTTTTCTGCGCGTAATCTGCTGTGCAACAAAAACCAACCGCTACCGCGGTGG 1099  
QY 5066 TTTTGTGTCGGGATCAAGAGCTACCACTCTTTTCCGAGGTAACTGGCTTCAGCAGAG 5125  
DB 1098 TTTTGTGTCGGGATCAAGAGCTACCACTCTTTTCCGAGGTAACTGGCTTCAGCAGAG 1039  
QY 5126 CGCAGATACCAAAATCTAGTCTCTTCTAGTGTAGCGTATGAGCCACCACTTCAAGAACT 5185  
DB 1038 CGCAGATACCAAAATCTAGTCTCTTCTAGTGTAGCGTATGAGCCACCACTTCAAGAACT 979  
QY 5186 CTGTAGCACCGCTTACATACCTCGCTCTGCTTAATCTGTTTACAGATGCGCTGCTGCCAGTG 5245  
DB 978 CTGTAGCACCGCTTACATACCTCGCTCTGCTTAATCTGTTTACAGATGCGCTGCTGCCAGTG 919  
QY 5246 GCGATAAGTCTGCTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGCGCGAGC 5305

DB 918 GCGATAAGTCTGCTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGCGCAGC 859  
QY 5306 GGTTCGGCTCTGACGCGGGGTTCTGTGACACAGCCAGCTTGGAGCGAAACGACCTTACACCG 5365  
DB 858 GGTTCGGCTCTGACGCGGGGTTCTGTGACACAGCCAGCTTGGAGCGAAACGACCTTACACCG 799  
QY 5366 AACTGAGATACCTTACAGCGGTGAGCTATGAGAAAGCGCAACGCTTCCGAAAGGAGAAAGG 5425  
DB 798 AACTGAGATACCTTACAGCGGTGAGCTATGAGAAAGCGCAACGCTTCCGAAAGGAGAAAGG 739  
QY 5426 CGACACAGATCTCGGTAAAGCGGCGAGGGTCCGAAACAGAGAGCGCACGAGGGAGCTTCCAG 5485  
DB 738 CGACACAGATCTCGGTAAAGCGGCGAGGGTCCGAAACAGAGAGCGCACGAGGGAGCTTCCAG 679  
QY 5486 GGGGAAACCGCTGCTGATCTTTATAGTCTCTGCGGGTTTCGCCACCTCTGACTTCAGCGCTC 5545  
DB 678 GGGGAAACCGCTGCTGATCTTTATAGTCTCTGCGGGTTTCGCCACCTCTGACTTCAGCGCTC 619  
QY 5546 GATTTTGTGATCTCGTCAAGGGGCGGAGCGCTATGGAATAAGCGCAACCGCGCT 5605  
DB 618 GATTTTGTGATCTCGTCAAGGGGCGGAGCGCTATGGAATAAGCGCAACCGCGCT 559  
QY 5606 TTTTACGGTTCTGCGCTTTGCTGCGCTTTTGTCTCACATGTTCTTCTCGCTGCTATCCC 5665  
DB 558 TTTTACGGTTCTGCGCTTTTGTCTGCGCTTTTGTCTCACATGTTCTTCTCGCTGCTATCCC 499  
QY 5666 CTGATTTCTGTGATAACCGCTATTAACCGCTTTTGTGAGTGAGTATACCGCTCGCGCAGCC 5725  
DB 498 CTGATTTCTGTGATAACCGCTATTAACCGCTTTTGTGAGTGAGTATACCGCTCGCGCAGCC 439  
QY 5726 GAACGACCGAGCGCAGCGAGTCAAGTGAAGCGAGGAGCGAGCGCCCAATACGCAAC 5785  
DB 438 GAACGACCGAGCGCAGCGAGTCAAGTGAAGCGAGGAGCGAGCGCCCAATACGCAAC 379  
QY 5786 CGCTCTCCCGCGCGCTTGGCGGATTCATTAATGCGAG 5822  
DB 378 CGCTCTCCCGCGCGCTTGGCGGATTCATTAATGCGAG 342

RESULT 8  
AAD09980/C  
ID AAD09980 standard; DNA; 5919 BP.  
XX  
AC AAD09980;  
XX  
DT 12-SBP-2001 (first entry)  
XX  
DE pHSP-GUS construct for transfection of pacific oysters.  
XX  
KW Fruitfly; fertility; reproduction; gametogenesis; microinjection; GUS;  
KW glucuronidase; infection; plasmid pHSP-GUS; HSP; heat shock promoter;  
KW embryogenesis; ds.  
XX  
OS Drosophila melanogaster.  
OS Unidentified.  
OS Chimeric.  
XX  
PN WO200148224-A1.  
XX  
XX 05-JUL-2001.  
XX  
PP 22-DEC-2000; 2000WO-AU001596.  
XX  
XX 24-DEC-1999; 99AU-00004884.  
XX  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
XX  
PI Thresher R, Hinds L, Hardy C, Whyard S, Vignarajan S, Grewe PM;  
PI Patil J;  
XX  
DR WPI; 2001-425672/45.  
XX

PT Novel construct for preventing embryogenesis in animals comprises native  
PT promoter, blocking DNA which abrogates function of crucial gene and  
PT genetic switch to regulate expression/repression of blocker/gene  
knockout.

XX Example 8; Page 196-198; 241pp; English.

CC The invention relates to a construct which allows animals to be bred in  
CC captivity but renders them infertile in the wild by allowing reversible  
CC control over fertility and reproduction. The construct comprises a native  
CC promoter, a blocking DNA sequence contoured for and designed to abrogate  
CC a crucial gene's function or to cause its mis-expression, and a genetic  
CC switch to regulate controlled expression/repression of the blocker/gene  
CC knockout. The construct is useful for preventing embryogenesis or  
CC gametogenesis in animals by stably transforming an animal cell with the  
CC construct by microinjection, transfection or infection, where the  
CC construct stably integrates into the genome by homologous recombination,  
CC and implanting the cell into a host organism, where a whole animal  
CC develops from the implanted cell. The present DNA sequence is plasmid  
CC construct pHPSP-GUS construct used for transfection of Pacific oysters.  
CC The plasmid contains glucuronidase (GUS) gene under the control of  
CC Drosophila heat shock promoter (dHSP)

XX Sequence 5919 BP; 1504 A; 1474 C; 1414 G; 1527 T; 0 U; 0 Other;

Query Match 46.0%; Score 2675.4; DB 4; Length 5919;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2676; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	3146	TATAGTGAAGTCTATTACAAATTCACCTGCGCTGCTTTTACAACTGCTGACTGGGAAAC	3205
DB	5919	TATAGTGAAGTCTATTACAAATTCACCTGCGCTGCTTTTACAACTGCTGACTGGGAAAC	5860
QY	3206	CCTGGCGTTACCCAACTTAATGCGCTTTCAGCAATCCCCCTTTGCGCAGCTGCGTAAT	3265
DB	5859	CCTGGCGTTACCCAACTTAATGCGCTTTCAGCAATCCCCCTTTGCGCAGCTGCGTAAT	5800
QY	3266	AGCGAAGGCGCGCGATGCGCTTCCCAACAGTTGCGCAGCTGGAATGGCGAATGG	3325
DB	5799	AGCGAAGGCGCGCGATGCGCTTCCCAACAGTTGCGCAGCTGGAATGGCGAATGG	5740
QY	3326	ACGCGCCTGTAGCGCGCATTAAGCGCGCGGTGTGGTGTACGCGCAGCTGACCG	3385
DB	5739	ACGCGCCTGTAGCGCGCATTAAGCGCGCGGTGTGGTGTACGCGCAGCTGACCG	5680
QY	3386	CTACACTTGCAGCGCCTAGCGCGCTCTCTTCCGCTTCTTCCCTTCTTCTCGCCA	3445
DB	5679	CTACACTTGCAGCGCCTAGCGCGCTCTCTTCCGCTTCTTCCCTTCTTCTCGCCA	5620
QY	3446	CGTTTCGCGCGCTTTCGCGCTCAAGCTCTAAATCGCGCGCTCCCTTTAGGGTTCCGATTTA	3505
DB	5619	CGTTTCGCGCGCTTTCGCGCTCAAGCTCTAAATCGCGCGCTCCCTTTAGGGTTCCGATTTA	5560
QY	3506	GAGCTTTACGCGCCTGACCGCGCAAACTTGAATTTGGGTGATGTTTACGTAAGTGGC	3565
DB	5559	GAGCTTTACGCGCCTGACCGCGCAAACTTGAATTTGGGTGATGTTTACGTAAGTGGC	5500
QY	3566	CATCGCCTCTGATAGACGGTTTTTTCGCGCTTTGAGCTTGCAGCTTCTTTAATAGTG	3625
DB	5499	CATCGCCTCTGATAGACGGTTTTTTCGCGCTTTGAGCTTGCAGCTTCTTTAATAGTG	5440
QY	3626	GACTCTTGTTCCTCAACTGGAAACAACCTCAACCTTATCTCGGCTCTTCTTTGATTTAT	3685
DB	5439	GACTCTTGTTCCTCAACTGGAAACAACCTTATCTCGGCTCTTCTTTGATTTAT	5380
QY	3686	AAGGGATTTTCGCGCTTATGCTTTTAAATAATGAGCTGATTTTAAACAAATTTTA	3745
DB	5379	AAGGGATTTTCGCGCTTATGCTTTTAAATAATGAGCTGATTTTAAACAAATTTTA	5320
QY	3746	ACGGGAATTTTAAACAAATTTAAGCTTTTACAAATTTGCGCTGATGCGGTATTTCTCCTT	3805
DB	5319	ACGGGAATTTTAAACAAATTTAAGCTTTTACAAATTTGCGCTGATGCGGTATTTCTCCTT	5260

QY	3806	ACGCATCTGTGCGCTATTTCACACCGCATACAGGTGGCACCTTTTCGGGGAATGTGCGG	3865
DB	5259	ACGCATCTGTGCGGTATTTCACACCGCATACAGGTGGCACCTTTTCGGGGAATGTGCGG	5200
QY	3866	GAACCCCTATTGTTTATTTCCTAAATATACATTCAAAATATGATATCGCTCATGACAAAT	3925
DB	5199	GAACCCCTATTGTTTATTTCCTAAATATACATTCAAAATATGATATCGCTCATGACAAAT	5140
QY	3926	AACCTGATTAATGCTTCAATAATATTGAAAAGAGAGATGAGTATTCAACATTTTCC	3985
DB	5139	AACCTGATTAATGCTTCAATAATATTGAAAAGAGAGATGAGTATTCAACATTTTCC	5080
QY	3986	GTGTGCGCTTATTTCCTTTTTCGGGCAATTTTCCTTCTGTTTTCCTCACCAGAAA	4045
DB	5079	GTGTGCGCTTATTTCCTTTTTCGGGCAATTTTCCTTCTGTTTTCCTCACCAGAAA	5020
QY	4046	CGCTGGTGAAGTAAAGATGCTGAAAGATCAAGTTGGGTGACAGAGTGGGTTATCAGAAC	4105
DB	5019	CGCTGGTGAAGTAAAGATGCTGAAAGATCAAGTTGGGTGACAGAGTGGGTTATCAGAAC	4960
QY	4106	TGATCTCAACAGCGGTGAAGATCTTGAAGTCTTTCGCGCCCGAAGAACGTTTTCATAGA	4165
DB	4959	TGATCTCAACAGCGGTGAAGATCTTGAAGTCTTTCGCGCCCGAAGAACGTTTTCATAGA	4900
QY	4166	TGACATCTTTTAAAGTTCTGCTATGTGCGCGGTATTATCCCGTATTTCACGCGCGGCAAG	4225
DB	4899	TGACATCTTTTAAAGTTCTGCTATGTGCGCGGTATTATCCCGTATTTCACGCGCGGCAAG	4840
QY	4226	AGCAACTCGGTGCGCGCATACACTATTCTCAGAAATGACTTGGTTGAGTACTCACCAGTCA	4285
DB	4839	AGCAACTCGGTGCGCGCATACACTATTCTCAGAAATGACTTGGTTGAGTACTCACCAGTCA	4780
QY	4286	CAGAAAAGCATCTTACCGGATGGCATGAGTAAGAGAAATTAATGCGTGTGCGCATACCA	4345
DB	4779	CAGAAAAGCATCTTACCGGATGGCATGAGTAAGAGAAATTAATGCGTGTGCGCATACCA	4720
QY	4346	TGATGATTAACACTCGCGCCAACTTCTTCTGACAAAGTCTGAGGACCGAAGGAGCTAA	4405
DB	4719	TGATGATTAACACTCGCGCCAACTTCTTCTGACAAAGTCTGAGGACCGAAGGAGCTAA	4660
QY	4406	CCGCTTTTTCACAAACATGCGGGGATCATGTAACTCGCTTGTGATGTTGGGAAACCGGAGC	4465
DB	4659	CCGCTTTTTCACAAACATGCGGGGATCATGTAACTCGCTTGTGATGTTGGGAAACCGGAGC	4600
QY	4466	TGAATGAAGCATACCAAAACGACGAGCGTGAACAACAAGTCTGTAGCAATGCGCAACA	4525
DB	4599	TGAATGAAGCATACCAAAACGACGAGCGTGAACAACAAGTCTGTAGCAATGCGCAACA	4540
QY	4526	CGTTTCGCAAACTTAATTAACCTGGCGAACTACTTACTCTAGCTTCCGCGCAACAATTAATAG	4585
DB	4539	CGTTTCGCAAACTTAATTAACCTGGCGAACTACTTACTCTAGCTTCCGCGCAACAATTAATAG	4480
QY	4586	ACTGGATGGAGCGCGATAAAGTTTCAGGACCACTTCTGCGCTCGGCGCTTCGCGCTGGCT	4645
DB	4479	ACTGGATGGAGCGCGATAAAGTTTCAGGACCACTTCTGCGCTCGGCGCTTCGCGCTGGCT	4420
QY	4646	GTTTATTGCTGATAAATCTGGAGCGGTGAGCGTGGTCTCGCGGTATCAATTGAGCAC	4705
DB	4419	GTTTATTGCTGATAAATCTGGAGCGGTGAGCGTGGTCTCGCGGTATCAATTGAGCAC	4360
QY	4706	TGGGCGCAGATGGTAAGCCCTCCGCTATCGTAGTTATCTACACGACGGGAGTCAAGGAA	4765
DB	4359	TGGGCGCAGATGGTAAGCCCTCCGCTATCGTAGTTATCTACACGACGGGAGTCAAGGAA	4300
QY	4766	CTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCTCTCACTGATTAAGCATTTGGT	4825
DB	4299	CTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCTCTCACTGATTAAGCATTTGGT	4240
QY	4826	AACCTGACAGCAAGTTTACTCATATATATCTTTAGATTGATTTTAAACCTTCATTTTAAAT	4885
DB	4239	AACCTGACAGCAAGTTTACTCATATATATCTTTAGATTGATTTTAAACCTTCATTTTAAAT	4180
QY	4886	TTAAAGGATCTAGTGAAGATCCTTTTGTGATAATCTCTCATGACCAAAATCCCTTAAACGTG	4945

Db 4179 TTTAAAGGATCTAGTGAAGATCTTTTGTATATCTCATGACCAAAATCCCTTAACGTG 4120  
Qy 4946 AGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAGAGATCTTCTTGAGATC 5005  
Db 4119 AGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAGAGATCTTCTTGAGATC 4060  
Qy 5006 CTTTTCGTTCCGCGTAAATCTGCTGCTTGTGCAAAACAAAACCAACCGCTACCGAGCGGTG 5065  
Db 4059 CTTTTCGTTCCGCGTAAATCTGCTGCTTGTGCAAAACAAAACCAACCGCTACCGAGCGGTG 4000  
Qy 5066 TTTGTTTCCGCGTCAAGAGCTACCACTCTTTTCCGAAGGTAACTGGCTTCAGCAGAG 5125  
Db 3999 TTTGTTTCCGCGTCAAGAGCTACCACTCTTTTCCGAAGGTAACTGGCTTCAGCAGAG 3940  
Qy 5126 CGCAGATACCAAAATCTGCTTCTAGTGTAGCGGTAGTGTAGGCCACCACTTCAAGAACT 5185  
Db 3939 CGCAGATACCAAAATCTGCTTCTAGTGTAGCGGTAGTGTAGGCCACCACTTCAAGAACT 3980  
Qy 5186 CTGTAGCAGCGCTACATACCTGCTCTGCTATCTCTGTTTACAGTGGCTGCTGCCAGTG 5245  
Db 3879 CTGTAGCAGCGCTACATACCTGCTCTGCTATCTCTGTTTACAGTGGCTGCTGCCAGTG 3820  
Qy 5246 GCGATAAGTCTGTCCTTACCGGTTGACTCAAGACGATAGTTACCGGATAAGCGCAGC 5305  
Db 3819 GCGATAAGTCTGTCCTTACCGGTTGACTCAAGACGATAGTTACCGGATAAGCGCAGC 3760  
Qy 5306 GGTCTGGCTGAACGGGGGTTCTGTCACACAGCCAGCTTGGAGCGAAACCACTTACACCG 5365  
Db 3759 GGTCTGGCTGAACGGGGGTTCTGTCACACAGCCAGCTTGGAGCGAAACCACTTACACCG 3700  
Qy 5366 AACTGAGATACCTACAGCGTGTATGAGAAAGCGCCAGCTTCCGAAGGAGAAAG 5425  
Db 3699 AACTGAGATACCTACAGCGTGTATGAGAAAGCGCCAGCTTCCGAAGGAGAAAG 3640  
Qy 5426 CGCACAGGTATCGGTAAAGCGGAGGTGCGAACGAGAGAGCGACGAGGAGCTTCCAG 5485  
Db 3639 CGCACAGGTATCGGTAAAGCGGAGGTGCGAACGAGAGAGCGACGAGGAGCTTCCAG 3580  
Qy 5486 GGGGAAACGCTGTGTATCTTTATAGTCTCTGCGGGTTTCGCCACCTCTGACTTGAGCGTC 5545  
Db 3579 GGGGAAACGCTGTGTATCTTTATAGTCTCTGCGGGTTTCGCCACCTCTGACTTGAGCGTC 3520  
Qy 5546 GATTTTGTGATGCTGTGAGGGGGGCGAGCGCTATGGAAGAAACGCCAGCAACCGCGCT 5605  
Db 3519 GATTTTGTGATGCTGTGAGGGGGGCGAGCGCTATGGAAGAAACGCCAGCAACCGCGCT 3460  
Qy 5606 TTTTACGTTCTCGGCTTTTGTGCTTTTGTCTCAGATGTTCTTCTGCGGTATCCC 5665  
Db 3459 TTTTACGTTCTCGGCTTTTGTGCTTTTGTCTCAGATGTTCTTCTGCGGTATCCC 3400  
Qy 5666 CTGATTTCTGTGATACCGGATTAACCGCTTTTGTGAGTGAGCTGATACGCTCGCGCAGCC 5725  
Db 3399 CTGATTTCTGTGATACCGGATTAACCGCTTTTGTGAGTGAGCTGATACGCTCGCGCAGCC 3340  
Qy 5726 GAACGACCGAGCGGAGTGTAGTGAGCGAGAGCGGAGAGCGCCGATACGCAAC 5785  
Db 3339 GAACGACCGAGCGGAGTGTAGTGAGCGAGAGCGGAGAGCGCCGATACGCAAC 3280  
Qy 5786 CGCTCTCCCGCGGTTGCGCGATTCATTAAATGCGAG 5822  
Db 3279 CGCTCTCCCGCGGTTGCGCGATTCATTAAATGCGAG 3243

RESULT 9  
ID ACC44700/c  
XX ACC44700 standard; DNA: 3357 BP.  
AC ACC44700;  
XX  
XX  
DT 29-MAY-2003 (first entry)  
XX  
DE Plasmid pGEMEasyNOS nucleotide sequence SEQ ID NO:96.

XX  
KW Chromosome-based platform; artificial chromosome; eukaryotic chromosome; att site; integrase; recombinase; ACes; gene therapy; transgenic animal; platform artificial chromosome expression system; gene; ds.  
XX Synthetic.  
XX  
PN WO200297059-A2.  
XX  
PD 05-DBC-2002.  
XX  
XX 30-MAY-2002; 2002WO-US017452.  
XX  
PR 30-MAY-2001; 2001US-0294758P.  
PR 21-MAR-2002; 2002US-0366891P.  
XX  
XX (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.  
XX  
XX Perkins E, Perez C, Lindenbaum M, Greene A, Leung J, Fleming B; Stewart S, Shellard J;  
PI  
XX MPI; 2003-140461/13.  
XX  
XX Novel eukaryotic chromosome comprising one or many att sites which permits site-directed integration in the presence of lambda-integrase, useful for site-specific recombination-directed integration of DNA of interest.  
XX  
XX Example 12; Page 225; 272pp; English.  
XX  
XX The present invention describes a eukaryotic chromosome (I) comprising one or several att sites, where an att site is heterologous to the chromosome, and permits site-directed integration in the presence of lambda-integrase. Also described: (1) a platform artificial chromosome expression system (ACes) (II) comprising several sites that participate in recombinase catalyzed recombination; and (2) a method (M1) for introducing a heterologous nucleic acid into a platform artificial chromosome. (I) can be used in gene therapy. (M1) is useful for introducing a heterologous nucleic acid molecule into a platform artificial chromosome, preferably an ACes. (II) is useful for producing a transgenic animal (e.g. a fish, insect, reptile, amphibian, arachnid, or mammal) by introducing (II) by cell fusion, lipid-mediated transfection, by a carrier system, microinjection, microcell fusion, electroporation, microprojectile bombardment or direct DNA transfer into an embryonic cell, preferably a stem cell or an embryo. (II) comprises a heterologous nucleic acid that encodes a therapeutic product which is useful for making a library of ACes comprising random portions of a genome. ACC44612 to ACC44732 and ABP96650 to ABP96657 represent sequences used in the exemplification of the present invention  
XX  
SQ Sequence 3357 BP; 854 A; 821 C; 851 G; 831 T; 0 U; 0 Other;  
Query Match 45.9%; Score 2670.2; DB 7; Length 3357;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 2698; Conservative 0; Mismatches 8; Indels 3; Gaps 2;  
Qy 3114 CGGCCCGCATGCGAGCTGCGGCCCAATTCGCCCTATAGTGTGCTGTATTACAAATTCATGG 3173  
Db 2988 CGGGAGCATCGAGCTGCGGCCCAATTCGCCCTATAGTGTGCTGTATTACAAATTCATGG 2929  
Qy 3174 CGGTGCTTTTACAACTGCTGCTGCGGAAACCTCGCGTTACCCAACTTAATCGCCTG 3233  
Db 2928 CGGTGCTTTTACAACTGCTGCTGCGGAAACCTCGCGTTACCCAACTTAATCGCCTG 2869  
Qy 3234 CAGCACAATCCCTTTTTCGCCAGCTGCGCTAATAGCAAGAGGCCCGCACGATCGCCCTT 3293  
Db 2868 CAGCACAATCCCTTTTTCGCCAGCTGCGCTAATAGCAAGAGGCCCGCACGATCGCCCTT 2809  
Qy 3294 CCACACAGTTGCGCAGCTGGAATGGCGAATGGAGCGCCCTGTAGCGGCCATTAAGCGC 3353  
Db 2808 CCACACAGTTGCGCAGCTGGAATGGCGAATGGAGCGCCCTGTAGCGGCCATTAAGCGC 2749  
Qy 3354 GCGCGGTGTGTGTTACGCGCAGCTGACCGCTACACTTTCGCCAGCGCCCTAGCGCGCCG 3413

Db 2748 |||||GGCGGTTGGTGGTTACGCGAGGGTACCGCTACACTTGGCAGCGCCCTAGCGCCCGC 2689  
Qy 3414 TCCCTTTCGCTTCTTCCCTCTCCCTTCTCGCCACAGTTTCGCGGCTTTCGCGCTCAAGCTCT 3473  
Db 2688 TCCCTTTCGCTTCTTCCCTCTCCCTTCTCGCCACAGTTTCGCGGCTTTCGCGCTCAAGCTCT 2629  
Qy 3474 AAATCGGGGGTCCCTTTAGGGTTCCGATTTAGAGCTTTACGGCACTTCGACCGCAAAAA 3533  
Db 2628 AAATCGGGGGTCCCTTTAGGGTTCCGATTTAGAGCTTTACGGCACTTCGACCGCAAAAA 2569  
Qy 3534 ACTTGATTTGGTGATGCTTACGTPAGTGGCCATCGCCCTGATAGAGCGTTTTTCGCC 3593  
Db 2568 ACTTGATTTAGGGTATGCTTACGTPAGTGGCCATCGCCCTGATAGAGCGTTTTTCGCC 2509  
Qy 3594 TTTGACGTTGGAGTCCAGTTCTTTAAATAGTGGACTCTTTGTTCCAAAATGGAAACAACACT 3653  
Db 2508 TTTGACGTTGGAGTCCAGTTCTTTAAATAGTGGACTCTTTGTTCCAAAATGGAAACAACACT 2449  
Qy 3654 CAACCTATCTCGGCTCTATCTTTTGAATTAAGGGATTTTTCGCGATTTGCGGCTATTG 3713  
Db 2448 CAACCTATCTCGGCTCTATCTTTTGAATTAAGGGATTTTTCGCGATTTGCGGCTATTG 2389  
Qy 3714 GTTAAATAATGAGCTGATTTTAAACAATATTTAACCGGAATTTTAAACAATATTTAACGCT 3773  
Db 2388 GTTAAATAATGAGCTGATTTTAAACAATATTTAACCGGAATTTTAAACAATATTTAACGCT 2329  
Qy 3774 TACAAATTTTCGCTGATCGGCTATTTTCTCTTACGCACTCTGTCGCTATTTTCAACCGCA 3833  
Db 2328 TACAAATTT - CCTGATCGGCTATTTTCTCTTACGCACTCTGTCGCTATTTTCAACCGCA 2271  
Qy 3834 TACAGTGGCACTTTTCGGGGAATTTGCGGGAAATTTGCGGGAAACCCCTATTTGTTTCTAAAT 3893  
Db 2270 T - CAGTGGGCACTTTTCGGGGAATTTGCGGGAAATTTGCGGGAAACCCCTATTTGTTTCTAAAT 2212  
Qy 3894 ACATTTCAATATGATCGCTCATGAGACAATAACCTGATAAATGCTTCAATATATG 3953  
Db 2211 ACATTTCAATATGATCGCTCATGAGACAATAACCTGATAAATGCTTCAATATATG 2152  
Qy 3954 AAAAAGGAAGATGATGATTTCAACATTTTCGCTGTCGCCCTTATTTCCCTTTTTCGCGC 4013  
Db 2151 AAAAAGGAAGATGATGATTTCAACATTTTCGCTGTCGCCCTTATTTCCCTTTTTCGCGC 2092  
Qy 4014 ATTTTGGCTTCTGTTTTCACCCAGAAACGCTGGTGAAGATTTAAAGATGCTGAAGA 4073  
Db 2091 ATTTTGGCTTCTGTTTTCACCCAGAAACGCTGGTGAAGATTTAAAGATGCTGAAGA 2032  
Qy 4074 TCAGTTGGTGCAGAGTGGTTTACATCGAATGATTTGATTTTAAAGTTTCTGCTATGTTG 4133  
Db 2031 TCAGTTGGTGCAGAGTGGTTTACATCGAATGATTTGATTTTAAAGTTTCTGCTATGTTG 1972  
Qy 4134 GAGTTTTCGCCCGGAAGACGTTTTCATGATGAGCACTTTTAAAGTTTCTGCTATGTTG 4193  
Db 1971 GAGTTTTCGCCCGGAAGACGTTTTCATGATGAGCACTTTTAAAGTTTCTGCTATGTTG 1912  
Qy 4194 CGGGTATTATCCGTTATGACGCGGGCAGAGCAACTCGGTCGCGCATACACTATTC 4253  
Db 1911 CGGGTATTATCCGTTATGACGCGGGCAGAGCAACTCGGTCGCGCATACACTATTC 1852  
Qy 4254 TCAGATGATGTTGGTTGATGATCTACACAGTTCACAGAAAGCATCTTACGGATGGCATGC 4313  
Db 1851 TCAGATGATGTTGGTTGATGATCTACACAGTTCACAGAAAGCATCTTACGGATGGCATGC 1792  
Qy 4314 AGTAAGAGAAATATGCACTGCTGCCATTAACCATGATGATTAACACTGCGGCCAACTTACT 4373  
Db 1791 AGTAAGAGAAATATGCACTGCTGCCATTAACCATGATGATTAACACTGCGGCCAACTTACT 1732  
Qy 4374 TCTGACAAACGATCGAGGACCGAAGAGCTAACCGCTTTTTCGACAAACATGCGGGATCA 4433  
Db 1731 TCTGACAAACGATCGAGGACCGAAGAGCTAACCGCTTTTTCGACAAACATGCGGGATCA 1672  
Qy 4434 TGTAACTCGCTTGTATCTGTTGGACCGGAGCTGAATGAACCAATACCAACGACGAGCG 4493

Db 1671 TGTAACTCGCTTGTATCTGTTGGBAACCGAGCTGAATGAAGCCATACCAACGACGAGCG 1612  
Qy 4494 TGACACCAACGATGCTGTAGCAATGGCAACAAACGTTGTCGCAAACTATTAACTGGCGAACT 4553  
Db 1611 TGACACCAACGATGCTGTAGCAATGGCAACAAACGTTGTCGCAAACTATTAACTGGCGAACT 1552  
Qy 4554 ACTTACTCTAGCTTCCCGGCAACAAATTAATAGATGGAATGGAGGGGATTAAGTTTGCAGG 4613  
Db 1551 ACTTACTCTAGCTTCCCGGCAACAAATTAATAGATGGAATGGAGGGGATTAAGTTTGCAGG 1492  
Qy 4614 ACCACTTCTCGCTCGGCGCTTTCGCGCTGCTGTTTATTTGCTGATAAATCTGGAGCCGG 4673  
Db 1491 ACCACTTCTCGCTCGGCGCTTTCGCGCTGCTGTTTATTTGCTGATAAATCTGGAGCCGG 1432  
Qy 4674 TGAGCGTGGTCTCGCGCTATCATTTGACGACTGCGGGCCAGATGTAAGCCCTCCGCTAT 4733  
Db 1431 TGAGCGTGGTCTCGCGCTATCATTTGACGACTGCGGGCCAGATGTAAGCCCTCCGCTAT 1372  
Qy 4734 GGTAGTTATCTACAGCAACCGGGAGTTCAGCAACTATGATGTAACGAAATAGACAGATCCG 4793  
Db 1371 GGTAGTTATCTACAGCAACCGGGAGTTCAGGCAACTATGATGTAACGAAATAGACAGATCCG 1312  
Qy 4794 TGAGATAGTGCCTCACTGATTTAAGCAATTTGCTGTAACCTGTCAGACCAAGTTTACTCATATAT 4853  
Db 1311 TGAGATAGTGCCTCACTGATTTAAGCAATTTGCTGTAACCTGTCAGACCAAGTTTACTCATATAT 1252  
Qy 4854 ACTTTAGATTTGATTTAAACCTTCATTTTAAATTTTAAAGGATCTAGGTGAAGATCTTTT 4913  
Db 1251 ACTTTAGATTTGATTTAAACCTTCATTTTAAATTTTAAAGGATCTAGGTGAAGATCTTTT 1192  
Qy 4914 TGATATCTCATGACCAAAATCCCTTAACGTGAGTTTTGCTTCCACTGAGCGTTCAGACCC 4973  
Db 1191 TGATATCTCATGACCAAAATCCCTTAACGTGAGTTTTGCTTCCACTGAGCGTTCAGACCC 1132  
Qy 4974 CGTAGAAAGATCAAGGATCTTCTTGAGATCCTTTTTTTCTGCGGTAATCTGCTGCTT 5033  
Db 1131 CGTAGAAAGATCAAGGATCTTCTTGAGATCCTTTTTTTCTGCGGTAATCTGCTGCTT 1072  
Qy 5034 GCAACAAAAAACAACCGCTACACGCGTGGTTTTGTTGCGGATCAAGAGTACCAAC 5093  
Db 1071 GCAACAAAAAACAACCGCTACACGCGTGGTTTTGTTGCGGATCAAGAGTACCAAC 1012  
Qy 5094 TCTTTTTCGAGGTTAACTGGCTTCAGCAGAGCGCAGATACCAATATCTGCTCTCTAGT 5153  
Db 1011 TCTTTTTCGAGGTTAACTGGCTTCAGCAGAGCGCAGATACCAATATCTGCTCTCTAGT 952  
Qy 5154 GTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCTCT 5213  
Db 951 GTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCTCT 892  
Qy 5214 GCTAATCTGTTACAGTGGCTGCTGCGAGTGAATGATGCTGCTTTTACCGGGTTGA 5273  
Db 891 GCTAATCTGTTTACAGTGGCTGCTGCGAGTGAATGATGCTGCTTTTACCGGGTTGA 832  
Qy 5274 CTCAGACGATGATTACCGGATTAAGCGCGGTCGGGCTGAACCGGGGGTTCTGTCAC 5333  
Db 831 CTCAGACGATGATTACCGGATTAAGCGCGGTCGGGCTGAACCGGGGGTTCTGTCAC 772  
Qy 5334 ACAGCCAGCTTTGAGCGCAACGACCTTACACCGAACTGAGATACCTTACAGCTGAGCTATG 5393  
Db 771 ACAGCCAGCTTTGAGCGCAACGACCTTACACCGAACTGAGATACCTTACAGCTGAGCTATG 712  
Qy 5394 AGAAGCGCCACGCTTCCGAAAGGGAGAAAGCGGACAGGATTCGGTAAGCGGCAAGGT 5453  
Db 711 AGAAGCGCCACGCTTCCGAAAGGGAGAAAGCGGACAGGATTCGGTAAGCGGCAAGGT 652  
Qy 5454 CGGAACGAGGACGACGAGGAGCTTTCAGGGGGAACCGCTGCTATCTTTTAGTCC 5513  
Db 651 CGGAACGAGGACGACGAGGAGCTTTCAGGGGGAACCGCTGCTATCTTTTAGTCC 592  
Qy 5514 TGTGGGGTTTCCCACTCTGACTTGAAGCTCGATTTTGTGATGCTCTGTCAGGGGGGG 5573  
Db 591 TGTGGGGTTTCCCACTCTGACTTGAAGCTCGATTTTGTGATGCTCTGTCAGGGGGGG 532



QY 5574 GAGCCTATGAAAAACCCAGCAACCGCGGCTTTTACGGTTCCTGCGCCTTTTCTGCTGGCC 5633  
DB 531 GAGCCTATGAAAAACCCAGCAACCGCGGCTTTTACGGTTCCTGCGCCTTTTCTGCTGGCC 472  
QY 5634 TTTTGTCTCACAATGTTCTTCTGCTGTTATCCCTGATTCTGTGATTAACCGTATTACCGC 5693  
DB 471 TTTTGTCTCACAATGTTCTTCTGCTGTTATCCCTGATTCTGTGATTAACCGTATTACCGC 412  
QY 5694 CTTTGTAGTGTAGTATACCGCTCGCGCAGCCGCAACGACGAGCGCAGCGAGTCAAGTGA 5753  
DB 411 CTTTGTAGTGTAGTATACCGCTCGCGCAGCCGCAACGACGAGCGCAGTCAAGTGA 352  
QY 5754 CGAGGAGCGGAGAGCGCCCAATACGCAACCGCTCTCCCGCGGCTTGGCGGATTCA 5813  
DB 351 CGAGGAGCGGAGAGCGCCCAATACGCAACCGCTCTCCCGCGGCTTGGCGGATTCA 292  
QY 5814 TTAATGCAG 5822  
DB 291 TTAATGCAG 283

## RESULT 10

ABT16597/c

ID ABT16597 standard; DNA; 3357 BP.

XX ABT16597;

DT 03-APR-2003 (first entry)

XX Artificial plant chromosome related plasmid DNA SEQ ID No 7.

XX Plant artificial chromosome; PAC; transgenic plant; vaccine;

KW blood factor; herbicide; stress; agronomical; nutrient quality;

KW bacterial artificial chromosome; BAC; yeast artificial chromosome; YAC;

KW ds.

XX Unidentified.

OS Unidentified.

XX WO200296923-A1.

XX 05-DEC-2002.

XX 30-MAY-2002; 2002WO-US017451.

XX 30-MAY-2001; 2001US-0294687P.

PR 04-JUN-2001; 2001US-0296329P.

XX (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.

XX (AGRI-) AGRISOMA INC.

XX Perez C, Fabijanski SF, Perkins E;

XX WPI; 2003-140436/13.

XX Producing artificial chromosome by introducing a nucleic acid into plant

PT cell, selecting artificial chromosome that has one or more repeat regions

PT with equivalent amounts of euchromatic and heterochromatic nucleic acids.

XX Example 5; Page 238; 269pp; English.

XX The invention relates to a novel method for producing plant artificial

XX chromosomes. The invention also relates to methods for targeting

CC insertion of heterologous DNA into plant artificial chromosomes, methods

CC for delivery of plant chromosomes to selected cells and tissues. The

CC isolated plant artificial chromosome (PAC) is useful for producing a

CC transgenic plant, which involves introducing the PAC into a plant cell.

CC The PAC comprises a heterologous nucleic acid encoding a gene product

CC such as enzymes, antisense RNA, rDNA, structural proteins, marker

CC proteins, ligands, receptors, ribozymes, therapeutic proteins, and

CC biopharmaceutical proteins, vaccines, blood factors, antigens, hormones,

CC cytokines, growth factors, antibodies, or a product that provides for

CC resistance to diseases, insects, herbicides, or stress in a plant. The

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC heterologous nucleic acid optionally encodes a product that provides an  
CC agronomically important trait in the plant, e.g. a product that alters  
CC nutrient use and/or improves the nutrient quality of the plant. The  
CC heterologous nucleic acid is contained within a bacterial artificial  
CC chromosome (BAC) or a yeast artificial chromosome (YAC). This  
CC polynucleotide sequence represents the DNA of a plasmid used in the  
CC method of the invention  
XX

SQ Sequence 3357 BP; 854 A; 821 C; 851 G; 831 T; 0 U; 0 Other;

Query Match 45.9%; Score 2670.2; DB 7; Length 3357;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 2698; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

QY 3114 CGCCCGATGCGAGCTGCGGCCCAATTCGCCCTATAGTGTGATTAACAATTCACGTG 3173

DB 2988 CGCGAGCATCGAGCTGCGGCCCAATTCGCCCTATAGTGTGATTAACAATTCACGTG 2929

QY 3174 CGCTCGTTTACCAACGTGCTGACTGGGAAACCCCTGGCGTTACCCAACTTAATCGCCTG 3233

DB 2928 CGCTCGTTTACCAACGTGCTGACTGGGAAACCCCTGGCGTTACCCAACTTAATCGCCTG 2869

QY 3234 CAGCACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGGCGCCGACCGATCGCCCTT 3293

DB 2868 CAGCACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGGCGCCGACCGATCGCCCTT 2809

QY 3294 CCCAACAGTTGCGCAGCTGGAATGGCGAATGAGCGCCCTGTAGCGGCGGATTAAGCGC 3353

DB 2808 CCCAACAGTTGCGCAGCTGGAATGGCGAATGAGCGCCCTGTAGCGGCGGATTAAGCGC 2749

QY 3354 GCGGGTGTGGTGTACCGCGAGCGTGACCTACACTTGCAGCGCCCTAGCGCCCGC 3413

DB 2748 GCGGGTGTGGTGTACCGCGAGCGTGACCTACACTTGCAGCGCCCTAGCGCCCGC 2689

QY 3414 TCCTTTCGCTTTCCTTCCCTTCTCGCCAGCTTTCGCCGCTTTCGCCGCTCAAGCTCT 3473

DB 2688 TCCTTTCGCTTTCCTTCCCTTCTCGCCAGCTTTCGCCGCTTTCGCCGCTCAAGCTCT 2629

QY 3474 AAATCGGGGCTCCCTTTTAGGGTTCGATTTAGAGCTTTACGGCAGCTTCGACGCCAATA 3533

DB 2628 AAATCGGGGCTCCCTTTTAGGGTTCGATTTAGAGCTTTACGGCAGCTTCGACGCCAATA 2569

QY 3534 ACTTGATTTGGGTGATGTTTACGTAGTGGGCCATCGCCCTGATAGACGCTTTTCGCC 3593

DB 2568 ACTTGATTTGGGTGATGTTTACGTAGTGGGCCATCGCCCTGATAGACGCTTTTCGCC 2509

QY 3594 TTTGAGCTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCTTTCCTTTCACACCT 3653

DB 2508 TTTGAGCTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCTTTCCTTTCACACCT 2449

QY 3654 CAACCTATCTCGCTCTATCTTTTGAATTAAGGGATTTTCGCCGATTTTCGCCCTATTG 3713

DB 2448 CAACCTATCTCGCTCTATCTTTTGAATTAAGGGATTTTCGCCGATTTTCGCCCTATTG 2389

QY 3714 GTTAAAAATGAGCTGATTTTAAACAAATATTTAAACGGAATTTTAAACAAATATTAACGTT 3773

DB 2388 GTTAAAAATGAGCTGATTTTAAACAAATATTTAAACGGAATTTTAAACAAATATTAACGTT 2329

QY 3774 TACAAATTTTCGCCCTGATGCGGTATTTTCTCCTTTCAGCATCTGTGCGGTATTTCACCGCA 3833

DB 2328 TACAAATTTTTCGCCCTGATGCGGTATTTTCTCCTTTCAGCATCTGTGCGGTATTTCACCGCA 2271

QY 3834 TACAGTGGCAGCTTTTCGCCGGAATTTTCGCCGGAACCCCTATTGTTTATTCTTCTAAAT 3893

DB 2270 T-CAGGTGGCAGCTTTTCGCCGGAATTTTCGCCGGAACCCCTATTGTTTATTCTTCTAAAT 2212

QY 3894 ACATTCAAAATATGATTCCTGCTCATGAGCAATAACCTCTGATAATGCTTCAATAATATTG 3953

DB 2211 ACATTCAAAATATGATTCCTGCTCATGAGCAATAACCTCTGATAATGCTTCAATAATATTG 2152

QY 3954 AAAAAAGGAGAGTATGAGTATTCAACATTTTCGCTGTGCGCCCTATTTCCTTTTTCGGCG 4013

DB 2151 AAAAAAGGAGAGTATGAGTATTCAACATTTTCGCTGTGCGCCCTATTTCCTTTTTCGGCG 2092



QY 4014 ATTTTGGCTTCCTGTTTTGCTCACCAGAAACGCTGGTGAAGTAAAGATGCTGAAGA 4073  
DB 2091 ATTTTGGCTTCCTGTTTTGCTCACCAGAAACGCTGGTGAAGTAAAGATGCTGAAGA 2032  
QY 4074 TCAGTTGGGTGCAAGATGGGTGTACATCGAACTGATCTCAACAGCGGTAAAGTCTCTTGA 4133  
DB 2031 TCAGTTGGGTGCAAGATGGGTGTACATCGAACTGATCTCAACAGCGGTAAAGTCTCTTGA 1972  
QY 4134 GAGTTTTGCGCCGGAAGAAAGTTTTTCCAATGATGAGCACTTTTAAAGTTCCTGCTATGCG 4193  
DB 1971 GAGTTTTGCGCCGGAAGAAAGTTTTTCCAATGATGAGCACTTTTAAAGTTCCTGCTATGCG 1912  
QY 4194 CGCGGTATTTATCCGTTATTTGACGCGGCAAGAGCAACTCGGTTCGCGCATACATATTC 4253  
DB 1911 CGCGGTATTTATCCGTTATTTGACGCGGCAAGAGCAACTCGGTTCGCGCATACATATTC 1852  
QY 4254 TCAGAAATGACTTGGTGTAGTACTACACAGTCAAGAAAGCACTTACGGATGGCATGAC 4313  
DB 1851 TCAGAAATGACTTGGTGTAGTACTACACAGTCAAGAAAGCACTTACGGATGGCATGAC 1792  
QY 4314 AGTAAGAGAAATTTATGACGTGCTGCCATAACCACTGATGATTAACACTGCGGCCAACTTACT 4373  
DB 1791 AGTAAGAGAAATTTATGACGTGCTGCCATAACCACTGATGATTAACACTGCGGCCAACTTACT 1732  
QY 4374 TCTGACAAACGATCGGAGGACCGAAGGAGCTAAACCGCTTTTTTGGCAACATGCGGGGATCA 4433  
DB 1731 TCTGACAAACGATCGGAGGACCGAAGGAGCTAAACCGCTTTTTTGGCAACATGCGGGGATCA 1672  
QY 4434 TGTAATCGCTTGTGATCTTGGGAACCGGAGCTGAATGAAGCCATACCAACGAGGAGCG 4493  
DB 1671 TGTAATCGCTTGTGATCTTGGGAACCGGAGCTGAATGAAGCCATACCAACGAGGAGCG 1612  
QY 4494 TGACACCAAGATGCTGTAGCAATGGCAACACGTTGCGCAACCTATTAACCTGGCGAAT 4553  
DB 1611 TGACACCAAGATGCTGTAGCAATGGCAACACGTTGCGCAACCTATTAACCTGGCGAAT 1552  
QY 4554 ACTTACTCTAGCTTCTCCGCGCAACAAATTAATAGACTGGATGGAGCGGATAAAGTTGCGAG 4613  
DB 1551 ACTTACTCTAGCTTCTCCGCGCAACAAATTAATAGACTGGATGGAGCGGATAAAGTTGCGAG 1492  
QY 4614 ACCACTTCTGGCTCGGCCCTTCGGCTGGCTGGTTTATGCTGATAAATCTGGAGCGG 4673  
DB 1491 ACCACTTCTGGCTCGGCCCTTCGGCTGGCTGGTTTATGCTGATAAATCTGGAGCGG 1432  
QY 4674 TGAGCGTGGGTCTCGCGGTATCATTTGACGACTGGGCGCAGATGGTAAGCCCTCCCGTAT 4733  
DB 1431 TGAGCGTGGGTCTCGCGGTATCATTTGACGACTGGGCGCAGATGGTAAGCCCTCCCGTAT 1372  
QY 4734 CGTATGTTATCTACAGCGGGAGTCAAGGCAACTATGGATGAACGAATAGACGATCGC 4793  
DB 1371 CGTATGTTATCTACAGCGGGAGTCAAGGCAACTATGGATGAACGAATAGACGATCGC 1312  
QY 4794 TGAGATAGGTGCTCCTCATGATTAGCATTTGCTGACCAAGTCTGACCAAGTCTCATATAT 4853  
DB 1311 TGAGATAGGTGCTCCTCATGATTAGCATTTGCTGACCAAGTCTGACCAAGTCTCATATAT 1252  
QY 4854 ACTTTAGATTGATTTAAATCTTCAATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTT 4913  
DB 1251 ACTTTAGATTGATTTAAATCTTCAATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTT 1192  
QY 4914 TGATAATCTCATGACCAAAATCCCTTAAACGTTGTTTCGTTCCATGAGCGGTGAGACCC 4973  
DB 1191 TGAATAATCTCATGACCAAAATCCCTTAAACGTTGTTTCGTTCCATGAGCGGTGAGACCC 1132  
QY 4974 CGTAGAAAAGATCAAGAGATCTTCTTGAGATCCTTTTTTCTGCGCGTAAATCTGCTGCTT 5033  
DB 1131 CGTAGAAAAGATCAAGAGATCTTCTTGAGATCCTTTTTTCTGCGCGTAAATCTGCTGCTT 1072  
QY 5034 GCAACCAAAAACCAACCGCTTACAGCGGTGCTTTGTTGTTGCGGATCAAGAGCTTACCAAC 5093  
DB 1071 GCAACCAAAAACCAACCGCTTACAGCGGTGCTTTGTTGTTGCGGATCAAGAGCTTACCAAC 1012

QY 5094 TCTTTTTCGGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAAATATCTGTCTCTTAGT 5153  
DB 1011 TCTTTTTCGGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAAATATCTGTCTCTTAGT 952  
QY 5154 GTAGCGGTAGTTAGGCGCACCACTTCAAGNACTCTCTAGCACCGCTTACATACCTCGCTCT 5213  
DB 951 GTAGCGGTAGTTAGGCGCACCACTTCAAGNACTCTCTAGCACCGCTTACATACCTCGCTCT 892  
QY 5214 GCTAATCTCTGTTTACAGTGGCTGCTGCCAGTGGCGATTAAGTCTGTCTTTACCGGGTGA 5273  
DB 891 GCTAATCTCTGTTTACAGTGGCTGCTGCCAGTGGCGATTAAGTCTGTCTTTACCGGGTGA 832  
QY 5274 CTCAGAGCATAGTTTACCGGATTAAGCGCAGCGCTCGGCTCGAAACGCGGGGTTCGTGCAC 5333  
DB 831 CTCAGAGCATAGTTTACCGGATTAAGCGCAGCGCTCGGCTCGAAACGCGGGGTTCGTGCAC 772  
QY 5334 ACAGCCAGCTTTGGAGCGAAACGACCTACACCGAAGCTGAGATACCTACAGCGTGAAGTATG 5393  
DB 771 ACAGCCAGCTTTGGAGCGAAACGACCTACACCGAAGCTGAGATACCTACAGCGTGAAGTATG 712  
QY 5394 AGAAGCGCCACGCTTCCGAAGGAGGAAAGCGGACAGGTATCCGTTAAGCGGACGGGT 5453  
DB 711 AGAAGCGCCACGCTTCCGAAGGAGGAAAGCGGACAGGTATCCGTTAAGCGGACGGGT 652  
QY 5454 CGBAACAGAGAGCGACAGGAGCTTCCAGGGGGGAAACGCTCGTATCTTTTATAGTCC 5513  
DB 651 CGBAACAGAGAGCGACAGGAGCTTCCAGGGGGGAAACGCTCGTATCTTTTATAGTCC 592  
QY 5514 TGTGCGGGTTTCGCCACCTCTGACCTTGGAGCGTCAATTTTGTGATGCTCGTCAGGGGGCG 5573  
DB 591 TGTGCGGGTTTCGCCACCTCTGACCTTGGAGCGTCAATTTTGTGATGCTCGTCAGGGGGCG 532  
QY 5574 GAGCTATGGAAGAAACGCGACCAACGCGGCTTTTACGGTTCTGCGCTTTTGTGGCC 5633  
DB 531 GAGCTATGGAAGAAACGCGACCAACGCGGCTTTTACGGTTCTGCGCTTTTGTGGCC 472  
QY 5634 TTTTGTCTCAGATGTTCTTCTGCTGCTGATTCCTGATTCCTGCTGATTAACCGCTTACCGC 5693  
DB 471 TTTTGTCTCAGATGTTCTTCTGCTGCTGATTCCTGATTCCTGCTGATTAACCGCTTACCGC 412  
QY 5694 CTTTGAGTGAGCTCATACCGCTTCGCGCAGCGAACCGAGCGCAGCGAGTCAAGTGAAG 5753  
DB 411 CTTTGAGTGAGCTCATACCGCTTCGCGCAGCGAACCGAGCGCAGCGAGTCAAGTGAAG 352  
QY 5754 CGAGGAAGCGGAAGAGCGCCCAATACGCAAAACCGCTCTCCCGCGCGTTCGCGCGATTC 5813  
DB 351 CGAGGAAGCGGAAGAGCGCCCAATACGCAAAACCGCTCTCCCGCGCGTTCGCGCGATTC 292  
QY 5814 TTAATGCGAG 5822  
DB 291 TTAATGCGAG 283

## RESULT 11

ADA94775/c  
ID ADA94775 standard; DNA; 4514 BP.

XX ADA94775;  
XX AC  
XX XX  
XX DT 20-NOV-2003 (first entry)  
XX XX  
XX DE Plasmid pGL2neo (+) #S8Q ID 4.  
XX XX  
XX KW Lymphocyte; vertebrate; antibody; genetically modified; ds.  
XX OS Synthetic.  
XX XX  
XX FN BP1321477-A1.  
XX XX  
XX PD 25-JUN-2003.  
XX XX  
XX PF 22-DEC-2001; 2001EP-00130805.  
XX XX

PR 22-DEC-2001; 2001EP-00130805.  
XX (GRAW/) GRAWUNDER U.  
PA (MELC/) MELCHERS G P.  
XX  
XX Grawunder U, Melchers GF;  
XX WPI; 2003-571360/54.  
XX  
XX Generating genetically modified vertebrate precursor lymphocytes for  
PT producing any heterologous antibody or binding protein comprises  
PT effecting differentiation of the precursor lymphocytes into mature  
PT lymphoid lineage cells.  
XX  
XX Example 5; Page 49-51; ilpp; English.  
XX  
XX The invention relates to a method for generating vertebrate lymphocytes  
CC that can be used for the production of any heterologous antibody, antigen  
CC receptor, artificial binding protein, or their functional fragments. The  
CC method of the invention comprises genetically modifying vertebrate  
CC precursor lymphocytes, and effecting differentiation of the precursor  
CC lymphocytes into mature lymphoid lineage cells either in vitro or in  
CC vivo. The method and the genetically modified and differentiated  
CC vertebrate lymphocytes are useful in the production of any heterologous  
CC antibody, artificial binding protein, antigen receptor, or their  
CC fragments, where the antibody is monoclonal or polyclonal, or partially  
CC resembles a human antibody, binding protein or antigen receptor. The  
CC antibodies are useful for the diagnosis, prevention and treatment of  
CC diseases. The method combines the advantages of both the phage display  
CC system (i.e. speed and flexibility in generating human antibodies, and  
CC the ability to modify and improve the properties of existing antibodies),  
CC and of the human immunoglobulin transgenic mouse technology (i.e. the  
CC ability to obtain high affinity antibodies due to affinity maturation  
CC occurring in the immune system, and the production of antibodies with  
CC physiologic and natural structural features). The current sequence  
CC represents the plasmid sequence pGL2neo (+). This sequence is used in an  
CC example from the invention as a template for the amplification of loxP  
CC site flanked neomycin.  
XX  
XX Sequence 4514 BP; 1108 A; 1184 C; 1171 G; 1051 T; 0 U; 0 Other;  
SQ  
Query Match 45.88; Score 2668.6; DB 8; Length 4514;  
Best Local Similarity 99.58; Pred. No. 0;  
Matches 2677; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
3132 GGCCCAATTCGCCCTATAGTAGTGTGATTAACAATTCACCTGGCGCGTGTTCACACGTC 3191  
4514 GGCCCAATTCGCCCTATAGTAGTGTGATTAACAATTCACCTGGCGCGTGTTCACACGTC 4455  
3192 GTGACTGGGAAACCCCTGGCGTTACCCCACTTAATGCGCTTGAGGACATCCCGCTTTCG 3251  
4454 GTGACTGGGAAACCCCTGGCGTTACCCCACTTAATGCGCTTGAGGACATCCCGCTTTCG 4395  
3252 CCAGCTGGCGTAATAGCGAAGAGCGCGCCGATCCGCTTCCCAACAGTTCGCGAGCC 3311  
4394 CCAGCTGGCGTAATAGCGAAGAGCGCGCCGATCCGCTTCCCAACAGTTCGCGAGCC 4335  
3312 TGAATGCGGAATGAGCGCGCCCTGTAGCGCGCATTAAGCGCGCGGTGTGTTGTTAC 3371  
4334 TGAATGCGGAATGAGCGCGCCCTGTAGCGCGCATTAAGCGCGCGGTGTGTTGTTAC 4275  
3372 GCGCAGGTGACCGCTACACTTGGCAGCGCCCTAGCGCGCGCTTTCGCTTCTTCCTCC 3431  
4274 GCGCAGGTGACCGCTACACTTGGCAGCGCCCTAGCGCGCGCTTTCGCTTCTTCCTCC 4215  
3432 TTCCTTTCTCGCCACAGTTCGCGCGCTTTCGCGCTCAAGCTCTAAATCGGGGGCTCCCTTT 3491  
4214 TTCCTTTCTCGCCACAGTTCGCGCGCTTTCGCGCTCAAGCTCTAAATCGGGGGCTCCCTTT 4155  
3492 AGGGTTCGGAATTTAGAGCTTTACGGCACTTCGACCGCAAAAACTTGATTTGGGTGATGG 3551  
4154 AGGGTTCGGAATTTAGAGCTTTACGGCACTTCGACCGCAAAAACTTGATTTGGGTGATGG 4095

QY 3552 TTCACTAGTGGGCATCGCCCTGATAGACGGTTTTTTCGCCCTTTTGAGCTTGGAGTCCAC 3611  
DB 4094 TTCACTAGTGGGCATCGCCCTGATAGACGGTTTTTTCGCCCTTTTGAGCTTGGAGTCCAC 4035  
QY 3612 GTTCTTTAATAGTGGACTCTTGTTCAAAATCGGAACAACTCAACCTATCTCGGTCTA 3671  
DB 4034 GTTCTTTAATAGTGGACTCTTGTTCAAAATCGGAACAACTCAACCTATCTCGGTCTA 3975  
QY 3672 TTCTTTTGAATTAAGGGAATTTTGGCGGATTTTGGCGCTATTTGGTTTAAAAAATAGAGTGA 3731  
DB 3974 TTCTTTTGAATTAAGGGAATTTTGGCGGATTTTGGCGGCTATTTGGTTTAAAAAATAGAGTGA 3915  
QY 3732 TTAACAAATATTTAAACCGGAATTTTAAACAAATATTTAAACGTTTACAAATTTTGGCCTGATGC 3791  
DB 3914 TTAACAAATATTTAAACCGGAATTTTAAACAAATATTTAAACGTTTACAAATTTTGGCCTGATGC 3855  
QY 3792 GGTATTTTCTCTTACGCACTCTGCGGTATTTTACACACCGCATACAGTGGCACTTTTCG 3851  
DB 3854 GGTATTTTCTCTTACGCACTCTGCGGTATTTTACACACCGCATACAGTGGCACTTTTCG 3795  
QY 3852 GGGAAATGTGCGCGGAACCCCTATTTTGTATTTTCTTAAATACATTTCAAAATATGTATCC 3911  
DB 3794 GGGAAATGTGCGCGGAACCCCTATTTTGTATTTTCTTAAATACATTTCAAAATATGTATCC 3735  
QY 3912 GCTCATGAGCAATTAACCCCTGATTAATGCTTCAATATATTTGAAAAGGAGATGATGAG 3971  
DB 3734 GCTCATGAGCAATTAACCCCTGATTAATGCTTCAATATATTTGAAAAGGAGATGATGAG 3675  
QY 3972 TATTTCAACATTTCCGTGTGCGCTTATTTCCCTTTTTCGCGCATTTTTCGCTTCTGTTTT 4031  
DB 3674 TATTTCAACATTTCCGTGTGCGCTTATTTCCCTTTTTCGCGCATTTTTCGCTTCTGTTTT 3615  
QY 4032 TGCTCAACCAAGAAACGCTGTGTAAGATTAAGATGCTGAAAGATCAGTTGGGTGCAAGT 4091  
DB 3614 TGCTCAACCAAGAAACGCTGTGTAAGATTAAGATGCTGAAAGATCAGTTGGGTGCAAGT 3555  
QY 4092 GGGTTACATCGAATCTCAACAGCGGTAGATCTTTGAGAGTTTTCGCCCGGAAGA 4151  
DB 3554 GGGTTACATCGAATCTCAACAGCGGTAGATCTTTGAGAGTTTTCGCCCGGAAGA 3495  
QY 4152 ACCTTTTCCAAATGATGAGCACTTTTAAAGTTTCTGCTATGTGGCGCGTATTTATCCCGTAT 4211  
DB 3494 ACCTTTTCCAAATGATGAGCACTTTTAAAGTTTCTGCTATGTGATACATTTATCCCGTAT 3435  
QY 4212 TGACGCGGGCAAGAGCAACTCGGTGCGCGCATACATTTCTCAGAAATGACTTGGTTGA 4271  
DB 3434 TGACGCGGGCAAGAGCAACTCGGTGCGCGCGCATACATTTCTCAGAAATGACTTGGTTGA 3375  
QY 4272 GTACTCACCAGTCAACAGAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTTATCAG 4331  
DB 3374 GTACTCACCAGTCAACAGAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTTATCAG 3315  
QY 4332 TGCTGCATTAACCATGAGTGAATAACACTTGGCGCAACTTCTTCTGCAACAGATCGGAGG 4391  
DB 3314 TGCTGCATTAACCATGAGTGAATAACACTTGGCGCAACTTCTTCTGCAACAGATCGGAGG 3255  
QY 4392 ACCGAAGAGCTAAACCGCTTTTTCGACAAACATGCGGGGATCATGTAACTGCGCTTGTATCG 4451  
DB 3254 ACCGAAGAGCTAAACCGCTTTTTCGACAAACATGCGGGGATCATGTAACTGCGCTTGTATCG 3195  
QY 4452 TTGGGAAACCGGAGCTCAATGAGCCCATACCAACAGAGCGGTGACACCAAGATGCGCTGT 4511  
DB 3194 TTGGGAAACCGGAGCTCAATGAGCCCATACCAACAGAGCGGTGACACCAAGATGCGCTGT 3135  
QY 4512 AGCAATGGCAACAACTGTGCGCAAACTATTAATCTGCGGAATCTTACTCTAGCTTCCCG 4571  
DB 3134 AGCAATGGCAACAACTGTGCGCAAACTATTAATCTGCGGAATCTTACTCTAGCTTCCCG 3075  
QY 4572 GCAACAAATTAATAGCTGAGTGGAGCGGATTAAGTTGAGGAGCACTTCTGCGCTCGGC 4631  
DB 3074 GCAACAAATTAATAGCTGAGTGGAGCGGATTAAGTTGAGGAGCACTTCTGCGCTCGGC 3015  
QY 4632 CCTTCGCGCTGGCTGTTTTATTTCTGTGATAAATCTGAGAGCGGTGAGGCTGCGCTGCGCG 4691







Db 424 TGGTTCAGGTAGTGGGCAATGACCTGATAGACGGTTTTTGGCCCTTTGACGTTGGAGTC 483  
Qy 3609 CACGTTCTTTAAATAGTGGACCTCTTGTGTTCCAAAATGGAACAACACTCAACCCCTATCTCGGT 3668  
Db 484 CACGTTCTTTAAATAGTGGACCTCTTGTGTTCCAAAATGGAACAACACTCAACCCCTATCTCGGT 543  
Qy 3669 CTATTCCTTTGANTTTAATAGGGAATTTGCGGATTTTGGCCCTATTTGGTTTAAANAATGAGCT 3728  
Db 544 CTATTCCTTTGANTTTAATAGGGAATTTTGGCCGATTTTGGCCCTATTTGGTTTAAANAATGAGCT 603  
Qy 3729 GATTTAAACAATATTTAAACGGAAATTTTAAACAATAATTAACGTTTAAACAATTTTCCGCTGA 3788  
Db 604 GATTTAAACAATATTTAAACGGAAATTTTAAACAATAATTAACGTTTAAACAATTTTCCGCTGA 661  
Qy 3789 TGGGTATTTTCTCTTACGCACTGTGCGGTATTTCAACCGGATATTCACCGGATACAGGTGCACTTT 3848  
Db 662 TGGGTATTTTCTCTTACGCACTGTGCGGTATTTTCAACCGGATATTCACCGGATACAGGTGCACTTT 720  
Qy 3849 TCGGGGAATGTGCGGGAACCCCTATTTGTTTATTTTCTTAAATACATTCAAATATGTA 3908  
Db 721 TCGGGGAATGTGCGGGAACCCCTATTTGTTTATTTTCTTAAATACATTCAAATATGTA 780  
Qy 3909 TCGGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTTGAAAAGGAAGATAT 3968  
Db 781 TCGGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTTGAAAAGGAAGATAT 840  
Qy 3969 GAGTATTTCAATTTCTGCTGCGGCTTATTTCCCTTTTTTGGCGCATTTTGCCTTCCCTGT 4028  
Db 841 GAGTATTTCAATTTCTGCTGCGGCTTATTTCCCTTTTTTGGCGCATTTTGCCTTCCCTGT 900  
Qy 4029 TTTTGTCTCACCCAGAACGCTGTGAAAGTAAAGATGCTGAAAGATCAGTTGGGTGCAAG 4088  
Db 901 TTTTGTCTCACCCAGAACGCTGTGAAAGTAAAGATGCTGAAAGATCAGTTGGGTGCAAG 960  
Qy 4089 AGTGGGTTTACATCGAACTGGATCTCAACACGGGTAAAGATCTTTGAGAGTTTTGCCCCGA 4148  
Db 961 AGTGGGTTTACATCGAACTGGATCTCAACACGGGTAAAGATCTTTGAGAGTTTTGCCCCGA 1020  
Qy 4149 AGAACGTTTTCGAATGATGAGCACTTTTAAAGTTCTGCTATGTTGGCGCGGTATTTATCCCG 4208  
Db 1021 AGAACGTTTTCGAATGATGAGCACTTTTAAAGTTCTGCTATGTTGGCGCGGTATTTATCCCG 1080  
Qy 4209 TATTGACGCGGCAAGAGCAACTCGTTCGCGCATACACTTCTCAGAACTACTTTGGT 4268  
Db 1081 TATTGACGCGGCAAGAGCAACTCGTTCGCGCATACACTTCTCAGAACTACTTTGGT 1140  
Qy 4269 TGAGTACTTCAACAGTCAAGAAAGCATCTTAACGATGGCATGACAGTAAAGAAATATG 4328  
Db 1141 TGAGTACTTCAACAGTCAAGAAAGCATCTTAACGATGGCATGACAGTAAAGAAATATG 1200  
Qy 4329 CAGTCTGCCATTAACCATGATGATTAACACTGCGGCGCAACTTACTTCTGACAAAGATGCG 4388  
Db 1201 CAGTCTGCCATTAACCATGATGATTAACACTGCGGCGCAACTTACTTCTGACAAAGATGCG 1260  
Qy 4389 AGGACCGAAGGAGCTAACCGCTTTTGGCAACATGCGGGATCATGTAACCTGCGCTGA 4448  
Db 1261 AGGACCGAAGGAGCTAACCGCTTTTGGCAACATGCGGGATCATGTAACCTGCGCTGA 1320  
Qy 4449 TCGTTGGGAACCGGAGCTGAATGAAGCCATTAACCAACGACGAGCGTGAACCAAGATGCC 4508  
Db 1321 TCGTTGGGAACCGGAGCTGAATGAAGCCATTAACCAACGACGAGCGTGAACCAAGATGCC 1380  
Qy 4509 TGTAGCAATGGCAACAAAGTTGGCGCAAACTATTAATCTGGCGAACTACTTCTAGGCTTC 4568  
Db 1381 TGTAGCAATGGCAACAAAGTTGGCGCAAACTATTAATCTGGCGAACTACTTCTAGGCTTC 1440  
Qy 4569 CCGGCAACAAATTAATAGCTGAGTGAAGGATTAAGTTGAGGACCACTTCGCGCTC 4628  
Db 1441 CCGGCAACAAATTAATAGCTGAGTGAAGGATTAAGTTGAGGACCACTTCGCGCTC 1500  
Qy 4629 GGCCTTCCGGCTCGGTGTTTATTTGCTGATAAATCTGGAGCGGTGAGCGTGGGTCTCG 4688

Db 1501 GGCCTTCCGGCTCGGTGTTTATTTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCG 1560  
Qy 4689 CGGTATCATTTGCAAGCACTGGGGCCAGATGGTAAGCCCTCCGTAATCGTAGTTATCTACAC 4748  
Db 1561 CGGTATCATTTGCAAGCACTGGGGCCAGATGGTAAGCCCTCCGTAATCGTAGTTATCTACAC 1620  
Qy 4749 GACGGGAGTCAAGCAACTATGATGAACGAATAGACAGATCGCTCAGATAGGTGCGCTC 4808  
Db 1621 GACGGGAGTCAAGCAACTATGATGAACGAATAGACAGATCGCTCAGATAGGTGCGCTC 1680  
Qy 4809 ACTGATTAAGCAATTTGGTAACTGTTCAGACCAAGTTTACTCATATATACCTTTAGATTTGATTT 4868  
Db 1681 ACTGATTAAGCAATTTGGTAACTGTTCAGACCAAGTTTACTCATATATACCTTTAGATTTGATTT 1740  
Qy 4869 AAAAATCTTATTTTAAATTTAAAAGGATCTAGGTGAAGATCTTTTTCATATAATCTCATGAC 4928  
Db 1741 AAAAATCTTATTTTAAATTTAAAAGGATCTAGGTGAAGATCTTTTTCATATAATCTCATGAC 1800  
Qy 4929 CAAAATCCCTTTAAACGTGAGTTTTTGGTTTCCACTGAGCGTCAAGCCCGTAGAAAAGATCAA 4988  
Db 1801 CAAAATCCCTTTAAACGTGAGTTTTTGGTTTCCACTGAGCGTCAAGCCCGTAGAAAAGATCAA 1860  
Qy 4989 AGGATCTTCTTGAGATCTCTTTTCTGCGCGTAATCTGCTGTGCTTGCACAAACAAAACC 5048  
Db 1861 AGGATCTTCTTGAGATCTCTTTTCTGCGCGTAATCTGCTGTGCTTGCACAAACAAAACC 1920  
Qy 5049 ACCGCTACAGCGGTGTTTGTGCGGATCAAGAGTACCAACTCTTTTTCGGAAGGT 5108  
Db 1921 ACCGCTACAGCGGTGTTTGTGCGGATCAAGAGTACCAACTCTTTTTCGGAAGGT 1980  
Qy 5109 AACTGGCTTCAGCAGAGCGCAGATACCAATACCTGCTTCTAGTGTAGCGTGTAGGTAGG 5168  
Db 1981 AACTGGCTTCAGCAGAGCGCAGATACCAATACCTGCTTCTAGTGTAGCGTGTAGGTAGG 2040  
Qy 5169 CCAACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTGCTCTGCTTAATCTCTGTACC 5228  
Db 2041 CCAACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTGCTCTGCTTAATCTCTGTACC 2100  
Qy 5229 AGTGGCTGTCCAGTGGCGATTAAGTCTGTCTTACCGGTTGGACTCAAGACGATGAT 5288  
Db 2101 AGTGGCTGTCCAGTGGCGATTAAGTCTGTCTTACCGGTTGGACTCAAGACGATGAT 2160  
Qy 5289 ACCGATTAAGGCGCAGCGTCCGGCTGAAACGGGGGTTTCGTGACACACAGCCAGCTTGA 5348  
Db 2161 ACCGATTAAGGCGCAGCGTCCGGCTGAAACGGGGGTTTCGTGACACACAGCCAGCTTGA 2220  
Qy 5349 GCGAAACGACTTACACCGAACTGAGATACCTTACAGCGTGAAGTATGAGAAAGCGCCAGCT 5408  
Db 2221 GCGAAACGACTTACACCGAACTGAGATACCTTACAGCGTGAAGTATGAGAAAGCGCCAGCT 2280  
Qy 5409 TCCGGAAGGAGAAAGCGCGACAGGTATCCGTTAAGCGGCGAGGTCGGAACAGGAGAGCG 5468  
Db 2281 TCCGGAAGGAGAAAGCGCGACAGGTATCCGTTAAGCGGCGAGGTCGGAACAGGAGAGCG 2340  
Qy 5469 CACGAGGAGCTTCCAGGGGGAACCGCTGCTATCTTTTATAGTCTCTGTCGGGTTTCGCCA 5528  
Db 2341 CACGAGGAGCTTCCAGGGGGAACCGCTGCTATCTTTTATAGTCTCTGTCGGGTTTCGCCA 2400  
Qy 5529 CTTCTGACTTGAGCGTTCGATTTTGTGATGCTGCTCAGGGGCGGAGCTCTATGGAAAAA 5588  
Db 2401 CTTCTGACTTGAGCGTTCGATTTTGTGATGCTGCTCAGGGGCGGAGCTCTATGGAAAAA 2460  
Qy 5589 CGGCAGCAACCGGCGCTTTTACGGTTCTGGCTTTTGGCTTTTGGCTTTTGGCTTTTGGCTTTT 5648  
Db 2461 CGGCAGCAACCGGCGCTTTTACGGTTCTGGCTTTTGGCTTTTGGCTTTTGGCTTTTGGCTTTT 2520  
Qy 5649 CTTTCTGCTGCTTATCCCTCTGATTTCTGCTGATTAACCGTATTTACCGCTTTTGGAGTGA 5708  
Db 2521 CTTTCTGCTGCTTATCCCTCTGATTTCTGCTGATTAACCGTATTTACCGCTTTTGGAGTGA 2580  
Qy 5709 TACCGCTCCCGCAGCCGAAACGACGCGAGCGAGTCAAGTCAAGGAGCGGAGAG 5768  
Db 2581 TACCGCTCCCGCAGCCGAAACGACGCGAGCGAGTCAAGTCAAGGAGCGGAGAG 2640



QY 5769 GCGCCCAATAGCAACCGGCTCTCCCGCGCGTGGCGGATTCATTAAATGCAG 5822  
DB 2641 GCGCCCAATAGCAACCGGCTCTCCCGCGCGTGGCGGATTCATTAAATGCAG 2694

RESULT 14  
ADA41728/c  
ID ADA41728 standard; DNA; 3404 BP.  
XX ADA41728;  
XX DT - 20-NOV-2003 (first entry)  
XX Plasmid pJO-20 DNA.  
DB RNA-dependent RNA polymerase; RdRP; plant virus; amplification system;  
KW ds; circular; cyclic.  
XX Synthetic.  
XX WO2003014366-A2.  
XX 20-FEB-2003.  
XX 29-JUL-2002; 2002WO-DE002863.  
XX 30-JUL-2001; 2001DE-01037444.  
XX (PROB-) PROBIOTEN AG.  
XX Sandig V, Jordan I;  
XX MPI; 2003-248302/24.  
XX  
XX Amplifying nucleic acid in animal cells, useful e.g. for gene therapy or  
PT vaccination, uses an RNA-dependent, RNA-polymerase of a plant virus.  
PT  
XX  
XX Example 2; Page 37-38; 39pp; German.  
XX  
XX This invention describes a novel method for amplifying nucleic acid in  
CC animal cells by introducing an RNA-dependent RNA polymerase (RdRP) and  
CC its associated promoters and cis-acting signals from a plant virus into  
CC the cells. RdRP is normally active in plant cells and the gene that  
CC encodes it can be recovered from such cells. Both the RdRP and the  
CC promoter are from plant viruses, particularly turnip crinkle virus and  
CC the amplified RNA is a modified satellite or genomic RNA of this virus.  
CC The method is particularly used for amplification of RNA (which may be  
CC mRNA for protein synthesis; an effector, e.g. antisense RNA or ribozyme,  
CC or genomic RNA) in animal cells, for (i) control of gene expression or  
CC (ii) for gene therapy or vaccination. When the system includes an  
CC inducible promoter, it permits strong and rapid expression of reporter  
CC genes in response to a test substance, especially where the promoter  
CC responds to the human immune deficiency virus or heavy metals, to produce  
CC a diagnostic system or biosensor, respectively. The method of the  
CC invention provides an inducible or constitutive, autonomous RNA-dependent  
CC RNA amplification system for animal cells that requires only one  
CC polymerase and does not use any viral structural genes or helper viruses.  
CC Amplification takes place in the cytoplasm without using any components  
CC potentially infectious for the host cells. Human 293 cells were  
CC transformed with (i) pJO-39, expressing a turnip crinkle virus 88 kD  
CC protein; (ii) pJO-60, expressed satellite RNA-C of the same virus in the  
CC sense orientation and a fusion of internal ribosome entry site and green  
CC fluorescent protein in the antisense orientation, and (iii) an expression  
CC vector for T7 RNA polymerase under control of the cytomegalovirus  
CC promoter. Expression of the reporter gene was detected by fluorescence  
CC microscopy. This sequence represents a fragment of plasmid, pJO-20 DNA  
CC used in the invention.  
XX  
XX Sequence 3404 BP; 858 A; 853 C; 863 G; 830 T; 0 U; 0 Other;

Query Match 45.48; Score 2641.4; DB 8; Length 3404;  
Best Local Similarity 99.74; Pred. No. 0;

		Matches 2668;	Conservative	0;	Mismatches	6;	Indels	3;	Gaps	2;
QY	3146	TATAGTCAGTCGTATTACAAATTCATCTGGCCGCTCGTTTACAAACGTCGTGACATGGGAAAC	3205							
DB	3404	TATAGTCAGTCGTATTACAAATTCATCTGGCCGCTCGTTTACAAACGTCGTGACATGGGAAAC	3345							
QY	3206	CCTGGCGTTACCCCAACTTAATTCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3265							
DB	3344	CCTGGCGTTACCCCAACTTAATTCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3285							
QY	3266	AGCGAAGAGGCCCGCAGTCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3325							
DB	3284	AGCGAAGAGGCCCGCAGTCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3225							
QY	3326	ACGCGCCCTGTAGCGCGCATTAAGCGCGCGGTGTGGTTACGCGAGCTGGCGTAAT	3385							
DB	3224	ACGCGCCCTGTAGCGCGCATTAAGCGCGCGGTGTGGTTACGCGAGCTGGCGTAAT	3165							
QY	3386	CTACACTTGCAGCGCCCTAGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3445							
DB	3164	CTACACTTGCAGCGCCCTAGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3105							
QY	3446	CGTTTCGCGCGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3505							
DB	3104	CGTTTCGCGCGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3045							
QY	3506	GAGCTTTACGCGCGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3565							
DB	3044	GAGCTTTACGCGCGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	2985							
QY	3566	CATCGCCCTGTAGCGCGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3625							
DB	2984	CATCGCCCTGTAGCGCGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	2925							
QY	3626	GACTCTTGTTCGCGCGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3685							
DB	2924	GACTCTTGTTCGCGCGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	2865							
QY	3686	AAGGGATTTTCGCGCGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3745							
DB	2864	AAGGGATTTTCGCGCGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	2805							
QY	3746	ACGCGAATTTTACAAATATTAAGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3805							
DB	2804	ACGCGAATTTTACAAATATTAAGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	2747							
QY	3806	ACGCGATCTGCGCGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3865							
DB	2746	ACGCGATCTGCGCGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	2688							
QY	3866	GAACCCCTATTTGTTTATTTTCTTAAATATTAAGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3925							
DB	2687	GAACCCCTATTTGTTTATTTTCTTAAATATTAAGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	2628							
QY	3926	AACCTGTATTAATGCTTCAATATTAATTAAGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	3985							
DB	2627	AACCTGTATTAATGCTTCAATATTAATTAAGCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	2568							
QY	3986	GTGTCGCCCTTATTTCCCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	4045							
DB	2567	GTGTCGCCCTTATTTCCCTTTCGCGCGCTTTCGAGCAGATCCCGCTTTCGCGAGCTGGCGTAAT	2508							
QY	4046	CGCTGTGAAAGTAAAGATGCTGAAAGATCAGTTGGGTGACGAGTGGGTACATCGAAC	4105							
DB	2507	CGCTGTGAAAGTAAAGATGCTGAAAGATCAGTTGGGTGACGAGTGGGTACATCGAAC	2448							
QY	4106	TGATCTCTACACGCGGTAAAGATCCTTGGAGAGTTTTCGCCCGCGAGAGAGCTTTTCCCAATGA	4165							
DB	2447	TGATCTCTACACGCGGTAAAGATCCTTGGAGAGTTTTCGCCCGCGAGAGAGCTTTTCCCAATGA	2388							
QY	4166	TGAGCACTTTTAAAGATTTCTGCTATGTGCGCGGTATTTATCCCGTATTTGACCGCGGCAAG	4225							
DB	2387	TGAGCACTTTTAAAGATTTCTGCTATGTGCGCGGTATTTATCCCGTATTTGACCGCGGCAAG	2328							



QY 4226 AGCAACTCGTCCCGCATACACTATTTCTCAGAAATGACTTGGTTGAGTACTCAACGATCA 4285  
DB 2327 AGCAACTCGTCCCGCATACACTATTTCTCAGAAATGACTTGGTTGAGTACTCAACGATCA 2268  
QY 4286 CAGAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTTATCAGTGTGCTGCCATACCA 4345  
DB 2267 CAGAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTTATCAGTGTGCTGCCATACCA 2208  
QY 4346 TGAGTGATAACACTCGCGGCAACTTTACTCTCGAACGATCGGAGGACCGAAGGAGCTAA 4405  
DB 2207 TGAGTGATAACACTCGCGGCAACTTTACTCTCGAACGATCGGAGGACCGAAGGAGCTAA 2148  
QY 4406 CGCTTTTGTGCAACAATGGGGGATCATGTAACTCGCTTTGATCTGTTGGGAAACCGGAGC 4465  
DB 2147 CGCTTTTGTGCAACAATGGGGGATCATGTAACTCGCTTTGATCTGTTGGGAAACCGGAGC 2088  
QY 4466 TGAATGAAGCCATACCAACGACGAGGTGACACACGATGCTGTAGCAATGGCAACAA 4525  
DB 2087 TGAATGAAGCCATACCAACGACGAGGTGACACACGATGCTGTAGCAATGGCAACAA 2028  
QY 4526 CGTTGGCGAAACTTAATACTGGCGAACTTACTCTAGCTTCCCGCAACAAATTAATAG 4585  
DB 2027 CGTTGGCGAAACTTAATACTGGCGAACTTACTCTAGCTTCCCGCAACAAATTAATAG 1968  
QY 4586 ACTGGATGGAGCGGATTAAGTTGCGAGGACCACTTCTCGCTCGGCCCTTCGCGCTGGCT 4645  
DB 1967 ACTGGATGGAGCGGATTAAGTTGCGAGGACCACTTCTCGCTCGGCCCTTCGCGCTGGCT 1908  
QY 4646 GGTTTATGCTGATAAATCTGGAGCGGTGAGCGTGGTCTCGCGTATCATTTGACGAC 4705  
DB 1907 GGTTTATGCTGATAAATCTGGAGCGGTGAGCGTGGTCTCGCGTATCATTTGACGAC 1848  
QY 4706 TGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACAGACGGGAGTCAGGCAA 4765  
DB 1847 TGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACAGACGGGAGTCAGGCAA 1788  
QY 4766 CTATGATGAACGAATAGACAGATCGTCGAGATAGTGGCTCACTGATTAAGCATTTGCT 4825  
DB 1787 CTATGATGAACGAATAGACAGATCGTCGAGATAGTGGCTCACTGATTAAGCATTTGCT 1728  
QY 4826 AACTGTCCAGACCAAGTTTACTCATATATCTTTAGATTGATTTAAACTTCATTTTAAAT 4885  
DB 1727 AACTGTCCAGACCAAGTTTACTCATATATCTTTAGATTGATTTAAACTTCATTTTAAAT 1668  
QY 4886 TTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTATGACCAAAATCCCTTAACGTG 4945  
DB 1667 TTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTATGACCAAAATCCCTTAACGTG 1608  
QY 4946 AGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAGGATCTTCTTGAGTNC 5005  
DB 1607 AGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAGGATCTTCTTGAGTNC 1548  
QY 5006 CTTTTTTTCTGCGCGTAATCTGCTGTGTCGTAACAAACCAACCAACCAACCAACCAACCA 5065  
DB 1547 CTTTTTTTCTGCGCGTAATCTGCTGTGTCGTAACAAACCAACCAACCAACCAACCAACCA 1488  
QY 5066 TTTGTTTGGCGATCAAGAGTACCAACTCTTTTTCGAAAGGTAACCTGGCTTCAGCAGAG 5125  
DB 1487 TTTGTTTGGCGATCAAGAGTACCAACTCTTTTTCGAAAGGTAACCTGGCTTCAGCAGAG 1428  
QY 5126 QGAGATACCAATATCTGCTCTAGTGTAGCGGTAGTGGCCACCACTTCAAGACT 5185  
DB 1427 QGAGATACCAATATCTGCTCTAGTGTAGCGGTAGTGGCCACCACTTCAAGACT 1368  
QY 5186 CTGTAGCACCGCTTACATACCTCGCTCTGCTATCTGTTTACAGTGGCTGCTGCCAGTG 5245  
DB 1367 CTGTAGCACCGCTTACATACCTCGCTCTGCTATCTGTTTACAGTGGCTGCTGCCAGTG 1308  
QY 5246 GCGATAGTCTGTCTTACCGGGTGGACTCAAGACGATAGTTACCGGATAGCGCGAGC 5305  
DB 1307 GCGATAGTCTGTCTTACCGGGTGGACTCAAGACGATAGTTACCGGATAGCGCGAGC 1248

QY 5306 GGTGCGGCTGAACCGGGGTTCTGTGACACAGCCAGCTTGGAGCGAACCTTACACCG 5365  
DB 1247 GGTGCGGCTGAACCGGGGTTCTGTGACACAGCCAGCTTGGAGCGAACCTTACACCG 1188  
QY 5366 AACTGAGATACCTTACAGCGTGAGCTATGAGAAAGCGCACGCTTCCGAAAGGAGAAAGG 5425  
DB 1187 AACTGAGATACCTTACAGCGTGAGCTATGAGAAAGCGCACGCTTCCGAAAGGAGAAAGG 1128  
QY 5426 QGACAGGTATCCGGTAAGCGGCGAGGTCGGAACAGAGAGCGCACGAGGAGCTTCCAG 5485  
DB 1127 QGACAGGTATCCGGTAAGCGGCGAGGTCGGAACAGAGAGCGCACGAGGAGCTTCCAG 1068  
QY 5486 QGGAACCGCTGTATCTTATAGTCTGTGCGGTTTGGCACCTCTGACTTGGAGGTC 5545  
DB 1067 QGGAACCGCTGTATCTTATAGTCTGTGCGGTTTGGCACCTCTGACTTGGAGGTC 1008  
QY 5546 GATTTTGTGATGCTCGTCAAGGGGCGGAGCCCTATGGAAGAAACCGCAGCAACCGGCT 5605  
DB 1007 GATTTTGTGATGCTCGTCAAGGGGCGGAGCCCTATGGAAGAAACCGCAGCAACCGGCT 948  
QY 5606 TTTTACGGTTCCTGGCTTTTGTGCTGGCTTTTGTCTCAGATGTTCTTCTGCGTTATCCC 5665  
DB 947 TTTTACGGTTCCTGGCTTTTGTGCTGGCTTTTGTCTCAGATGTTCTTCTGCGTTATCCC 888  
QY 5666 CTGATTTCTGGATTAACCGTATTTACCGCTTTTGTGAGTGAGCTATACCGCTCGCGAGCC 5725  
DB 887 CTGATTTCTGGATTAACCGTATTTACCGCTTTTGTGAGTGAGCTATACCGCTCGCGAGCC 828  
QY 5726 GAAACGACCGAGCGCAGCGAGTCACTGAGCGAGGAGGAGCGGAGCGCCCAATACGCAAC 5785  
DB 827 GAAACGACCGAGCGCAGCGAGTCACTGAGCGAGGAGGAGCGGAGCGCCCAATACGCAAC 768  
QY 5786 CGCTCTCCCGCGCTTGGCGGATTCATTAATGCGAG 5822  
DB 767 CGCTCTCCCGCGCTTGGCGGATTCATTAATGCGAG 731

RESULT 15  
AAD09981/c  
ID AAD09981 standard; DNA; 3968 BP.  
XX  
XX AAD09981;  
AC  
XX  
DT 12-SEP-2001 (first entry)  
XX  
DE pHP70-1MCS construct for transfection of pacific oysters.  
XX  
XX Fruitfly; fertility; reproduction; gametogenesis; microinjection; HSP;  
KW infection; plasmid pHP70-1MCS; heat shock promoter; embryogenesis; da.  
XX  
OS Drosophila melanogaster.  
OS Unidentified.  
OS Chimeric.  
XX  
XX WO200148224-A1.  
XX  
PD 05-JUL-2001.  
XX  
XX 22-DEC-2000; 2000WO-AU001596.  
XX  
XX 24-DEC-1999; 99AU-00004884.  
XX  
XX (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
XX  
PI Thresher R, Hinds L, Hardy C, Whyard S, Vignarajan S, Grewe PM;  
PI Patil J;  
XX  
XX WPI; 2001-425672/45.  
XX  
PT Novel construct for preventing embryogenesis in animals comprises native  
PT promoter, blocking DNA which abrogates function of crucial gene and  
PT genetic switch to regulate expression/repression of blocker/gene  
PT knockout.

XX	Example 9; Page 198-200; 241pp; English.	
PS	The invention relates to a construct which allows animals to be bred in	
XX	captivity but renders them infertile in the wild by allowing reversible	
CC	control over fertility and reproduction. The construct comprises a native	
CC	promoter, a blocking DNA sequence contoured for and designed to abrogate	
CC	a crucial gene's function or to cause its mis-expression, and a genetic	
CC	switch to regulate controlled expression/repression of the blocker/gene	
CC	knockout. The construct is useful for preventing embryogenesis or	
CC	gametogenesis in animals by stably transforming an animal cell with the	
CC	construct by microinjection, transfection or infection, where the	
CC	construct stably integrates into the genome by homologous recombination,	
CC	and implanting the cell into a host organism, where a whole animal	
CC	develops from the implanted cell. The present DNA sequence is plasmid	
CC	construct pHPSP70-1MCS used for transfection of Pacific oysters. The	
CC	plasmid contains Drosophila heat shock promoter 70 (dHSP70) and its poly	
CC	adenylation signal	
XX	Sequence 3968 BP; 1075 A; 917 C; 943 G; 1033 T; 0 U; 0 Other;	
QY	Query Match	
DB	Best Local Similarity 100.0%; Pred. No. 0;	
QY	Matched 2635; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
DB	3187 ACCTGCTGACTGGGAAAACCCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCC 3246	
QY	3187 ACCTGCTGACTGGGAAAACCCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCC 3909	
DB	3247 TTTTCGCAGCTGGGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGGC 3306	
QY	3247 TTTTCGCAGCTGGGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGGC 3849	
DB	3307 CAGCCTCAATGGCGAATGGAGCGCCCTGTAGCGCGCATTAAGCGGGCGGGTGTGGTG 3366	
QY	3307 CAGCCTCAATGGCGAATGGAGCGCCCTGTAGCGCGCATTAAGCGGGCGGGTGTGGTG 3789	
DB	3367 GTTAACGCGAGCGTGAACGCTGACACTTGCAGCGCCCTAGCGCCGCTCTCTTGCCTTTC 3426	
QY	3367 GTTAACGCGAGCGTGAACGCTGACACTTGCAGCGCCCTAGCGCCGCTCTCTTGCCTTTC 3729	
DB	3427 TTTCCCTTCTTTCGCGCAGCTTTCGCGCGCTTCCCGCTCAAGCTCTAAATCGGGGCTC 3486	
QY	3427 TTTCCCTTCTTTCGCGCAGCTTTCGCGCGCTTCCCGCTCAAGCTCTAAATCGGGGCTC 3669	
DB	3487 CTTTACGGGTTCCGATTTAGAGCTTTACGCGCACCTCGACCGCAAAAACCTTGATTGGGT 3546	
QY	3487 CTTTACGGGTTCCGATTTAGAGCTTTACGCGCACCTCGACCGCAAAAACCTTGATTGGGT 3609	
DB	3547 GATGGTTTCACTAGTGGGCCATCGCCCTGATAGACGGTTTTCGCCCTTTTGAACGTTGGAG 3606	
QY	3547 GATGGTTTCACTAGTGGGCCATCGCCCTGATAGACGGTTTTCGCCCTTTTGAACGTTGGAG 3549	
DB	3607 TCCAGGTTCTTAATAGTGAAGCTTGTTCGCAACTGGACACACCTCAACCTTATCTCG 3666	
QY	3607 TCCAGGTTCTTAATAGTGAAGCTTGTTCGCAACTGGACACACCTCAACCTTATCTCG 3489	
DB	3667 GTCTATTCTTTTCAATTAAGGATTTTTCGGGATTTTCGGGCTTATGGTTTAAAAAATGAG 3726	
QY	3667 GTCTATTCTTTTCAATTAAGGATTTTTCGGGATTTTCGGGCTTATGGTTTAAAAAATGAG 3429	
DB	3727 CTGATTTTAAACAAATATTAACGGGAATTTTAAACAAATATTAACGGTTTAAACATTTGCGCT 3786	
QY	3727 CTGATTTTAAACAAATATTAACGGGAATTTTAAACAAATATTAACGGTTTAAACATTTGCGCT 3369	
DB	3787 GATCGGATTTTCTTCAAGCATCTGTGGGATTTTCAACCGCATACAGGTGGCACT 3846	
QY	3787 GATCGGATTTTCTTCAAGCATCTGTGGGATTTTCAACCGCATACAGGTGGCACT 3309	
DB	3847 TTTTCGGGAAATGTCGCGGAACCCCTATTGTTGTTATTTTCTAAATACATTTCAAATATG 3906	
QY	3847 TTTTCGGGAAATGTCGCGGAACCCCTATTGTTGTTATTTTCTAAATACATTTCAAATATG 3249	
DB	3907 TATCGCTCATGAGACATTAACCTGTGATAAATGCTTCAATATATTTGAAAAAGGAGACT 3966	
QY	3907 TATCGCTCATGAGACATTAACCTGTGATAAATGCTTCAATATATTTGAAAAAGGAGACT 3189	
DB	3967 ATGAGTATTCAACATTTCCGTTGCGCCCTTATTCCTTTTTCGGGCAATTTTGCCTTCT 4026	
QY	3967 ATGAGTATTCAACATTTCCGTTGCGCCCTTATTCCTTTTTCGGGCAATTTTGCCTTCT 3129	
DB	4027 GTTTTTGCTCAACCCAGAAACCTCGTGAAGATAAAGATGCTGAAGATCAGTTGGGTGCA 4086	
QY	4027 GTTTTTGCTCAACCCAGAAACCTCGTGAAGATAAAGATGCTGAAGATCAGTTGGGTGCA 3069	
DB	4087 CGAGTGGTTACATCGAAGTCTCAACAGCGTGAAGATCCTTTCGAGGTTTTCGCCCC 4146	
QY	4087 CGAGTGGTTACATCGAAGTCTCAACAGCGTGAAGATCCTTTCGAGGTTTTCGCCCC 3009	
DB	4147 GAAGAACTTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATGTCGCGCGGTATATCC 4206	
QY	4147 GAAGAACTTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATGTCGCGCGGTATATCC 2949	
DB	4207 CGTATTGACCGCGGCAAGAGCAACTCGTCCGCGCATACACTATTTCTGAGAAATGACTTG 4266	
QY	4207 CGTATTGACCGCGGCAAGAGCAACTCGTCCGCGCATACACTATTTCTGAGAAATGACTTG 2889	
DB	4267 GTTTCAGTACTCACCAGTCAACAGAAAGCATCTTACGGATGCGATGACAGTAAGAAATTA 4326	
QY	4267 GTTTCAGTACTCACCAGTCAACAGAAAGCATCTTACGGATGCGATGACAGTAAGAAATTA 2829	
DB	4327 TGCAGTCTGTCATAAACCATGAGTGAATAACCTCGCGCAACTTCTGACAAACGATC 4386	
QY	4327 TGCAGTCTGTCATAAACCATGAGTGAATAACCTCGCGCAACTTCTGACAAACGATC 2769	
DB	4387 GGAGGACCGAAGAGTAAACCGCTTTTTCGCAACATGCGGGGATCATGTAACTCGCCTT 4446	
QY	4387 GGAGGACCGAAGAGTAAACCGCTTTTTCGCAACATGCGGGGATCATGTAACTCGCCTT 2709	
DB	4447 GATCGTTGGGAAACCGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACACGATG 4506	
QY	4447 GATCGTTGGGAAACCGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACACGATG 2649	
DB	4507 CCTGTAGCAATGCGCAACAACTTCGCGCAAACTTAACTGCGCAACTTACTCTAGCT 4566	
QY	4507 CCTGTAGCAATGCGCAACAACTTCGCGCAAACTTAACTGCGCAACTTACTCTAGCT 2589	
DB	4567 TCCCGGCAACAAATTAATAGACTGATGAGCGCGATTAAGTTGACGAGCACCTTCTGCGC 4626	
QY	4567 TCCCGGCAACAAATTAATAGACTGATGAGCGCGATTAAGTTGACGAGCACCTTCTGCGC 2529	
DB	4627 TCGGCCCTTCGCGCTGCGTGTATTTGCTGATTAATCTGAGCGCGTGAGCGTGGGTCT 4686	
QY	4627 TCGGCCCTTCGCGCTGCGTGTATTTGCTGATTAATCTGAGCGCGTGAGCGTGGGTCT 2469	
DB	4687 CGCGGTATCATTTGACGCACTGGGGCCAGATGATGAAGCCCTCCCGTATCGTAGTTATCTAC 4746	
QY	4687 CGCGGTATCATTTGACGCACTGGGGCCAGATGATGAAGCCCTCCCGTATCGTAGTTATCTAC 2409	
DB	4747 ACAGCGGGAGTCAGGCAACTATGATGAACGAAATAGACAGATCGCTGAGATAGGTGCC 4806	
QY	4747 ACAGCGGGAGTCAGGCAACTATGATGAACGAAATAGACAGATCGCTGAGATAGGTGCC 2349	
DB	4807 TCACGTGATTAAGCATTTGGTAAGTCTGACGCAAGTTTACTCATATATATCTTTAGATTGAT 4866	
QY	4807 TCACGTGATTAAGCATTTGGTAAGTCTGACGCAAGTTTACTCATATATATCTTTAGATTGAT 2289	
DB	4867 TTAATACTTCAATTTTAAATTTTAAAGGATCTAGGTGAAGATCTTTTGTGATAATCTCATG 4926	
QY	4867 TTAATACTTCAATTTTAAATTTTAAAGGATCTAGGTGAAGATCTTTTGTGATAATCTCATG 2229	
DB	4927 ACCAAATTCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGGTAGAAAGATC 4986	
QY	4927 ACCAAATTCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGGTAGAAAGATC 2169	
DB	4987 AAGGATCTCTTGAGATCTCTTTTCTGCGCGTAACTCTGCTGCTTCAACAAAAA 5046	



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 05:20:42 ; Search time 13309 Seconds  
(without alignments)  
13063.161 Million cell updates/sec

Title: US-09-924-197-1  
Perfect score: 5822  
Sequence: 1 ctggcagcaggtttccg.....tgccgattcattaatgag 5822

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*  
1: em\_estba.\*  
2: em\_esthba.\*  
3: em\_estmu.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_htc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_htc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_eston.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_hum.\*  
19: em\_gss\_hum.\*  
20: em\_gss\_hum.\*  
21: em\_gss\_hum.\*  
22: em\_gss\_hum.\*  
23: em\_gss\_hum.\*  
24: em\_gss\_hum.\*  
25: em\_gss\_hum.\*  
26: em\_gss\_hum.\*  
27: em\_gss\_hum.\*  
28: gb\_gss1.\*  
29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	1055.8	18.1	1070	9	AJ281552	AJ281552 4AJA-P6P1
2	966.4	16.9	1013	12	BM438846	BM438846 Iplvtr0015
C 3	924	15.9	1004	9	AJ281480	AJ281480 4AJA-P4G8
4	916.2	15.7	973	14	CD458281	CD458281 Fg08_08h0

5	895.6	15.2	917	14	CD458286	CD458286 Fg08_08a0
6	873	15.0	935	12	BG838279	BG838279 Gc01_10e0
7	865.4	14.9	918	14	CD459092	CD459092 Fg08_08e0
8	841	14.4	841	9	AL042026	AL042026 DKF2P434E
9	832.8	14.3	872	14	CD459085	CD459085 Fg08_08d0
C 10	825	14.2	1073	14	CF269652	CF269652 Fg08_08f0
11	824.6	14.2	854	12	BM438950	BM438950 Iplvtr0049
12	819.6	14.1	870	14	CD458333	CD458333 Fg08_08e0
C 13	815.6	14.0	1249	28	BZ572284	BZ572284 msh2_2572
14	814.4	14.0	1126	28	BZ577702	BZ577702 msh2_5533
15	808	13.9	819	14	CD649375	CD649375 CvGnd0008
C 16	807.6	13.9	965	28	BZ570738	BZ570738 msh2_1513
17	801.6	13.8	1483	28	BZ575896	BZ575896 msh2_4876
18	797.6	13.7	1011	28	BZ576726	BZ576726 msh2_5071
C 19	785.4	13.5	1574	28	BZ572566	BZ572566 msh2_2693
C 20	777.8	13.4	800	9	AJ281449	AJ281449 4AJA-P4D5
C 21	777.6	13.4	910	14	CD649387	CD649387 CvGnd0009
22	774.8	13.3	1336	28	BZ575810	BZ575810 msh2_4637
23	774.4	13.3	789	14	CD280920	CD280920 G44224_42
C 24	772.4	13.3	954	9	AL044364	AL044364 DKF2P434C
C 25	771.8	13.3	1067	9	AU081137	AU081137 AU081137
C 26	770.4	13.2	1370	28	BZ571721	BZ571721 msh2_2025
27	763.2	13.1	780	13	BQ825693	BQ825693 103012980
28	757	13.0	759	14	CD279661	CD279661 G43818_35
29	755.4	13.0	1463	28	BZ571475	BZ571475 msh2_1906
30	753.6	12.9	797	12	BM410796	BM410796 EST585323
31	745.4	12.8	786	14	CD458721	CD458721 Fg08_04e0
C 32	745.4	12.8	844	28	BZ574513	BZ574513 msh2_3706
C 33	743.6	12.8	832	12	BG923768	BG923768 602825893
34	741	12.7	741	14	CD279174	CD279174 G44221_83
35	741	12.7	966	28	BZ575002	BZ575002 msh2_4255
C 36	740.4	12.7	1089	9	AU081124	AU081124 AU081124
37	739.8	12.7	1003	28	BZ576686	BZ576686 msh2_5053
38	739.8	12.7	1112	28	BZ577534	BZ577534 msh2_5453
C 39	739.2	12.7	914	28	BZ569398	BZ569398 pac62-164
C 40	738.6	12.7	863	14	CF752100	CF752100 TGD89 Hum
41	733.8	12.6	950	28	BZ571129	BZ571129 msh2_1741
42	733.2	12.6	811	29	ATH517156	AJ571156 Arabidops
43	732.4	12.6	998	28	BZ576702	BZ576702 msh2_5060
44	732.4	12.6	1376	28	BZ571741	BZ571741 msh2_2034
45	730	12.5	730	14	CD281097	CD281097 G44224_38

ALIGNMENTS

RESULT 1	AJ281552/c	AJ281552	1070 bp	mRNA	linear	EST 30-JUN-2000
LOCUS	4AJA-P6P11-P	gambiae cDNA clone 4AJA-P6P11	1070 bp	mRNA	linear	EST 30-JUN-2000
DEFINITION	gambiae cDNA clone 4AJA-P6P11	gambiae cDNA clone 4AJA-P6P11	1070 bp	mRNA	linear	EST 30-JUN-2000
ACCESSION	AJ281552	AJ281552	1070 bp	mRNA	linear	EST 30-JUN-2000
VERSION	AJ281552.1	AJ281552.1	1070 bp	mRNA	linear	EST 30-JUN-2000
KEYWORDS	EST	EST	1070 bp	mRNA	linear	EST 30-JUN-2000
SOURCE	Anopheles gambiae	Anopheles gambiae	1070 bp	mRNA	linear	EST 30-JUN-2000
ORGANISM	Anopheles gambiae	Anopheles gambiae	1070 bp	mRNA	linear	EST 30-JUN-2000
REFERENCE	1 (bases 1 to 1070)	1 (bases 1 to 1070)	1070 bp	mRNA	linear	EST 30-JUN-2000
AUTHORS	Dinopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M.B. and Kafatos, F.C.	Dinopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M.B. and Kafatos, F.C.	1070 bp	mRNA	linear	EST 30-JUN-2000
TITLE	Anopheles gambiae pilot gene discovery project: Identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines	Anopheles gambiae pilot gene discovery project: Identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines	1070 bp	mRNA	linear	EST 30-JUN-2000
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	Proc. Natl. Acad. Sci. U.S.A.	1070 bp	mRNA	linear	EST 30-JUN-2000
MEDLINE	20300950	20300950	1070 bp	mRNA	linear	EST 30-JUN-2000
PUBMED	10841561	10841561	1070 bp	mRNA	linear	EST 30-JUN-2000
COMMENT	Contact: Dimopoulos G Fotis C. Kafatos laboratory European Molecular Biology Laboratory Meyerohofstrasse 1, 69117 Heidelberg, Germany.	Contact: Dimopoulos G Fotis C. Kafatos laboratory European Molecular Biology Laboratory Meyerohofstrasse 1, 69117 Heidelberg, Germany.	1070 bp	mRNA	linear	EST 30-JUN-2000



Qy 4132 GAGAGTTTTCGCCCGAAGAACGTTTTCCTCAATGATGAGCACTTTTAAAGTTCTGCTATGT 4191  
Dy 121 GAGAGTTTTCGCCCGAAGAACGTTTTCCTCAATGATGAGCACTTTTAAAGTTCTGCTATGT 180  
Qy 4192 GCGCGGTATTATCCCGATTATGACGCGCGGGAAGAGCAACTCGGTCCGCGATACACTAT 4251  
Dy 181 GCGCGGTATTATCCCGATTATGACGCGCGGGAAGAGCAACTCGGTCCGCGATACACTAT 240  
Qy 4252 TCTCAGATGACTTGGTTGAGTACTCACCAGTCAAGAAAGCACTTTACGGATGGCATG 4311  
Dy 241 TCTCAGATGACTTGGTTGAGTACTCACCAGTCAAGAAAGCACTTTACGGATGGCATG 300  
Qy 4312 ACAGTAAGAGATTATGACGTGCTGCCATAACCATGAGTATACACTTCGGGCCAACTTA 4371  
Dy 301 ACAGTAAGAGATTATGACGTGCTGCCATAACCATGAGTATACACTTCGGGCCAACTTA 360  
Qy 4372 CTTCTGACAAAGATCGGAGGACGAGGAGTAAACCGCTTTTTCGACAAATGCGGGAT 4431  
Dy 361 CTTCTGACAAAGATCGGAGGACGAGGAGTAAACCGCTTTTTCGACAAATGCGGGAT 419  
Qy 4432 C-ATGTAACTCGCTTGTATCGTTGGAAACCGGAGCTGAATGAAGCATACCAACGACGA 4490  
Dy 420 CAATGTAACTCGCTTGTATCGTTGGAAACCGGAGCTGAATGAAGCATACCAACGACGA 479  
Qy 4491 GCGTGACACCAAGTCTGTAGCAATGCGCAACCAACGTTGGCGCAAACTTAACTTGGCGA 4550  
Dy 480 GCGTGACACCAAGTCTGTAGCAATGCGCAACCAACGTTGGCGCAAACTTAACTTGGCGA 539  
Qy 4551 ACTACTTACTTACTGCTTCCCGGCAACATTAATAGACTGGATGGGCGGATTAAGTTGC 4610  
Dy 540 ACTACTTACTTACTGCTTCCCGGCAACATTAATAGACTGGATGGGCGGATTAAGTTGC 599  
Qy 4611 AGGACCACTTCTGCGCTCGGCGCTTCGCGCTGGCTGTTTATGCTGATTAATCTGAGC 4670  
Dy 600 AGGACCACTTCTGCGCTCGGCGCTTCGCGCTGGCTGTTTATGCTGATTAATCTGAGC 659  
Qy 4671 CGGTGAGGTGGGTCTCGCGGTATCATTTGACGCACTCGGGCCAGATGTAAGCCCTCCCG 4730  
Dy 660 CGGTGAGGTGGGTCTCGCGGTATCATTTGACGCACTCGGGCCAGATGTAAGCCCTCCCG 719  
Qy 4731 TATCGTAGTATCTACACGCGGGGATCAGGCACTATGATGAACCAATAGACAGAT 4790  
Dy 720 TATCGTAGTATCTACACGCGGGGATCAGGCACTATGATGAACCAATAGACAGAT 779  
Qy 4791 CGGTGAGTAGGTGCTCCTCACTGATTAAAGCATTTGTAAGTAACTGTCAGACCAAGTTTACTCAT 4850  
Dy 780 CGGTGAGTAGGTGCTCCTCACTGATTAAAGCATTTGTAAGTAACTGTCAGACCAAGTTTACTCAT 839  
Qy 4851 TATACTTTAGATTGATTTAAACCTTCATTTTAAATTTTAAAGGATCTAGGTGAAGATCCT 4910  
Dy 840 TATACTTTAGATTGATTTAAACCTTCATTTTAAATTTTAAAGGATCTAGGTGAAGATCCT 899  
Qy 4911 TTTTGATATCTCATGACCAAAATCCCTTACGTTAGTGGTTTCTTCCACTGAGGTGCTAGA 4970  
Dy 900 TTTTGATATCTCATGACCAAAATCCCTTAAAGTGGTTTCTTCCACTGAGGTGCTAGA 959  
Qy 4971 CCCCGTAGAAAGATCAAGGATCTTCTGAGATCCTTTTTCGCGCGTAAT 5024  
Dy 960 CCCCGTAGAAAGATCAAGGATCTTCTGAGATCCTTTTTCGCGCGTAAT 1013

RESULT 3  
AJ281480/c  
LOCUS  
DEFINITION 4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles  
gambiae cDNA clone 4A3A-P4G8, mRNA sequence.  
ACCESSION  
VERSION AJ281480  
KEYWORDS AJ281480.1 GI:6929360  
SOURCE Anopheles gambiae (African malaria mosquito)  
ORGANISM Anopheles gambiae  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;  
Anopheles.

## REFERENCE

1 (bases 1 to 1004)  
Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,  
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B.  
and Kafatos, F.C.  
Anopheles gambiae pilot gene discovery project: identification of  
mosquito innate immunity genes from expressed sequence tags  
generated from immune-competent cell lines  
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

10841561

Contact: Dimopoulos G

Fotis C. Kafatos Laboratory

European Molecular Biology Laboratory

Meyerothstrasse 1, 69117 Heidelberg, Germany.

Location/Qualifiers

## FEATURES

## source

1..1004

/organism="Anopheles gambiae"

/mol\_type="mRNA"

/strain="4A r/z"

/db\_xref="taxon:7165"

/clone="4A3A-P4G8"

/cell\_line="immune competent 4A3A"

/lab\_host="E. coli DH10B"

/clone\_lib="Anopheles gambiae immune competent 4A3A"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site 1: EcoRI; Site 2: NotI; sequenced from

forward priming site which reads from the 3' end of the

cDNA. The 4A3A is a directionally cloned and normalized

cDNA library that was constructed from the 4A3A cell line

oligo-T primed cDNA according to: Bonaldo, Lennon & Soares

(1996) : Normalization and Subtraction: Two approaches To

Facilitate Gene Discovery, Genome Research 6, 791-806."

## ORIGIN

Query Match 15.9%; Score 924; DB 9; Length 1004;

Best Local Similarity 99.0%; Pred. No. 5.9e-223;

Matches 1001; Conservative 1; Mismatches 2; Indels 7; Gaps 7;

Qy 4616 CACTTCTCGGCTCGGCCCTTCGGCTGGCTGTTTATTGCTGATAAATCTGAGCGCGTG 4675  
Dy 1004 CACTTCTCGGCTCGG-CCTTCGGCTGGCTGTTTATTGCTGATAAATCTGAGCGG-GTG 948  
Qy 4676 AGCGTGGTCTCGGCTATCATTTGACGCACTGGGCGGAGATGGTAAGCCCTCCCGTATCG 4735  
Dy 947 AGCGT-GGCTCGGCTATCATTTGACGCACT-GGCGGAGATGGTAAGCCCTCCCGTATCG 890  
Qy 4736 TAGTTATCTACACGCGGGAGTCAGGCACTATGATGAACCAATAGACAGATCGCTG 4795  
Dy 889 TAG-TATCTACACGCGGGAGTCAGGCACTATGATGAACCAATAGACAGATCGCTG 831  
Qy 4796 AGATAGTGGCTCACTGATTAAAGCATTTGTAAGTAACTGCTGAGCAAGTTTACTCATATATAC 4855  
Dy 830 AGATA-GTGCTCACTGATTAAAGCATTTGTAAGTAACTGCTGAGCAAGTTTACTCATATATAC 772  
Qy 4856 TTTAGATTGATTAAAGCATTTTAAAGTAACTGCTGAGCAAGTTTACTCATATATATAC 4915  
Dy 771 TTTAGATTGATTAAAGCATTTTAAAGTAACTGCTGAGCAAGTTTACTCATATATATAC 712  
Qy 4916 ATAATCTCATGACCAAAATCCCTTAAAGTAACTGCTGAGCAAGTTTACTCATATATATAC 4975  
Dy 711 ATAATCTCATGACCAAAATCCCTTAAAGTAACTGCTGAGCAAGTTTACTCATATATATAC 652  
Qy 4976 TAGAAAGATCAAGGATCTTCTGAGATCCTTTTCTGCGCTGTAATCTGCTGCTGC 5035  
Dy 651 TAGAAAGATCAAGGATCTTCTGAGATCCTTTTCTGCGCTGTAATCTGCTGCTGC 592  
Qy 5036 AAACAAAAAACCCCGCTACCGCGGTGTTTGTTCGCGATCAAGAGCTTACCAACTC 5095  
Dy 591 AAACAAAAAACCCCGCTACCGCGGTGTTTGTTCGCGATCAAGAGCTTACCAACTC 532  
Qy 5096 TTTTTCGAGGTAACTGGCTTTCAGCAGCGGAGTACCAATACTGCTCTTAGTGT 5155

```
Db 531 TTTTTCGAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAACTACTGTTCTTCTAGTGT 472
Qy 5156 AGCGGTAGTATAGGCGCACCACCTTCAAGAACTCTGTAGCAGCGCCTACATACCTCGCTCTGC 5215
Db 471 AGCGGTAGTATAGGCGCACCACCTTCAAGAACTCTGTAGCAGCGCCTACATACCTCGCTCTGC 412
Qy 5216 TAATCCTGTTACCAAGTGGCTGCTGCGCAGTGGCGATAAGTCGTGCTTACCGGGTTGGAAT 5275
Db 411 TAATCCTGTTACCAAGTGGCTGCTGCGCAGTGGCGATAAGTCGTGCTTACCGGGTTGGAAT 352
Qy 5276 CAAGCAGATAGTACCGGATAAGCGGAGCGGCTGGGCTGGAACGGGGGTTGCTGCGACAC 5335
Db 351 CAAGCAGATAGTACCGGATAAGCGGAGCGGCTGGGCTGGAACGGGGGTTGCTGCGACAC 292
Qy 5336 AGCCGAGCTTCGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAG 5395
Db 291 AGCCGAGCTTCGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAG 232
Qy 5396 AAAGCGGCACGCTTCCCGAAGGAGAGGCGGAGCGATATCCGGTAAGCGGCGGCTGCG 5455
Db 231 AAAGCGGCACGCTTCCCGAAGGAGAGGCGGAGCGATATCCGGTAAGCGGCGGCTGCG 172
Qy 5456 GAACGAGAGCGGCAAGCGGAGCTTCCAGGCGGAGAAAGCGGCTGATCTTTATAGTCCTG 5515
Db 171 GAACGAGAGCGGCAAGCGGAGCTTCCAGGCGGAGAAAGCGGCTGATCTTTATAGTCCTG 112
Qy 5516 TCGGTTTTCGCCACTCTGACTGAGCGTCCGATTTTGTGATGCTCGTCAGCGGGCGGA 5575
Db 111 TCGGTTTTCGCCACTCTGACTGAGCGTCCGATTTTGTGATGCTCGTCAGCGGGCGGA 52
Qy 5576 GCCTATGGAAGAAAGCGGAGCGGCTTTTACGGTTCTGCGGCTTTT 5626
Db 51 GCCTATGGAAGAAAGCGGAGCGGCTTTTACGGTTCTGCGGCTTTT 1
```

## RESULT 4

```
CD458281
LOCUS
DEFINITION
  Fg08_08h09 R Fg08 AAFPC ECORC Fusarium graminearum complex_substrate
  Gibberella zeae cDNA clone Fg08_08h09, mRNA sequence.
ACCESSION
  CD458281
VERSION
  CD458281.1 GI:31373021
KEYWORDS
  EST.
ORGANISM
  Gibberella zeae
  Gibberella zeae
  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
  Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE
  1 (bases 1 to 973)
  Watson,R.J., Heyes,R., Chapados,J., Couroux,P., Harris,L.J.,
  Hattori,J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A.,
  Sprott,D. and Tinker,N.A.
  A cDNA library prepared from Fusarium graminearum grown on a
  complex plant substrate
  Unpublished (2003)
JOURNAL
  Contact: Watson, Robert.J.
COMMENT
  Eastern Cereal and Oilseed Research Centre
  Agriculture and Agri-food Canada
  Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
  CANADA
  Tel: (613) 759-1655
  Fax: (613) 759-1701
  Email: watsonrj@agr.gc.ca.
FEATURES
  source
    1..973
    /organism="Gibberella zeae"
    /mol_type="mRNA"
    /strain="DAOM 180378"
    /db_xref="taxon:5518"
    /clone="Fg08_08h09"
    /tissue type="Mycelium"
    /dev stage="Asexual"
    /lab_host="E. coli DH10B"
```

```
/clone_lib="Fg08_AAFPC_ECORC_Fusarium_graminearum_complex_s
ubstrate"
/notes="Vector: pBluescript II+; Site 1: EcoRI; Site 2:
XhoI; Fusarium graminearum grown on a complex plant
substrate-- wheat leaves treated to remove most of the low
molecular weight, water-soluble components."

ORIGIN
Query Match 15.7%; Score 916.2; DB 14; Length 973;
Best Local Similarity 98.0%; Pred. No. 5.6e-221;
Matches 932; Conservative 5; Mismatches 13; Indels 1; Gaps 1;
Qy 3833 ATACAGGTGGCACTTTTCGGGAAATGTCGGGAAACCCCTATTGTTTATTTCTTAA 3892
Db 24 ATTTAGGTGGCASTTTTCGGGAAATGTCGGGAAACCCCTATTGTTTATTTCTTAA 83
Qy 3893 TACATTCAAAATATGATCCGCTCATGAGCAATAACCCCTGATAAATGCTTCAATAATATT 3952
Db 84 TACATTCAAAATATGATCCGCTCATGAGCAATAACCCCTGATAAATGCTTCAATAATATT 143
Qy 3953 GAAAAAGGAAGATATGAGTATTCAAATTTCCGCTGTCGCCCTTATCCCTTTTTCGGG 4012
Db 144 GAAAAAGGAAGATATGAGTATTCAAATTTCCGCTGTCGCCCTTATCCCTTTTTCGGG 203
Qy 4013 CATTTTGGCTTCTCTGTTTTCGCTCACCCAGAAACGCTGGTGAAGCTAAAGTGTCTGAG 4072
Db 204 CATTTTGGCTTCTCTGTTTTCGCTCACCCAGAGACGCTGGTGAAGTAAAGTGTCTGAG 263
Qy 4073 ATCAGTTGGGTGACAGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAGAGCTCTTG 4132
Db 264 ATCAGTTGGGTGACAGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAGAGCTCTTG 323
Qy 4133 AGAGTTTTCGCCCGGAAGAACGTTTTCATATGATGAGCACTTTTAAAGTTCTGCTATGTG 4192
Db 324 AGAGTTTTCGCCCGGAAGAACGTTTTCATATGATGAGCACTTTTAAAGTTCTGCTATGTG 383
Qy 4193 CGCGGCTATTATCCGCTATTGACGCGGGCAAGAGCAACTCGTCCGCCATACACTATT 4252
Db 384 CGCGGCTATTATCCGCTATTGACGCGGGCAAGAGCAACTCGTCCGCCATACACTATT 443
Qy 4253 CTCAAGATGACTGCTGTTGAGTACTCACCACTACAGAAAGCATCTTACGAGTGCATGA 4312
Db 444 CTCAAGATGACTGCTGTTGAGTACTCACCACTACAGAAAGCATCTTACGAGTGCATGA 503
Qy 4313 CAGTAAGAGAATTATGCAGTGTCTGCCATAACCATGAGTGATTAACACTCGGCGCAACTTAC 4372
Db 504 CAGTAAGAGAATTATGCAGTGTCTGCCATAACCATGAGTGATTAACACTCGGCGCAACTTAC 563
Qy 4373 TTCTGACACAGTCCGAGGACCGAAGGAGCTTAAACCGCTTTTTCGACACATGCGGGATC 4432
Db 564 TTCTGACACAGTCCGAGGACCGAAGGAGCTTAAACCGCTTTTTCGACACATGCGGGATC 623
Qy 4433 ATGTAACTCGCCTTTGATGCTTTGGGAAACCGGAGCTGAATGAAGCCATACCAACGAGC 4492
Db 624 ATGTAACTCGCCTTTGATGCTTTGGGAAACCGGAGCTGAATGAAGCCATACCAACGAGC 683
Qy 4493 GTGACACACAGTCCGCTGTAGCAATGCGCAACAACTTTCGCGCAAACTATTAACTGCGGAC 4552
Db 684 GTGACACACAGTCCGCTGTAGCAATGCGCAACAACTTTCGCGCAAACTATTAACTGCGGAC 743
Qy 4553 TACTTACTCTAGCTTCCGCGCAACAAATTAATAGACTGGATGAGCGGATTAAGTTGAG 4612
Db 744 TACTTACTCTAGCTTCCGCGCAACAAATTAATAGACTGGATGAGCGGATTAAGTTGAG 803
Qy 4613 GACCACCTTCTGCGCTCGGCCCTTCCGGCTGGCTGTTTATTTGCTGATAAATCTGGAGCG 4672
Db 804 GACCACCTTCTGCGCTCGGCCCTTCCGGCTGGCTGTTTATTTGCTGATAAATCTGGAGCG 863
Qy 4673 GTGAGCTGTGGTCTCGCGGTATTCATTTGAGCACTTGGGGCGAGATGATGAGCTCCCGTA 4732
Db 864 GTGAGCTGTGGTCTCGCGGTATTCATTTGAGCACTTGGGGCGAGATGATGAGAG-CTCCCGTA 922
Qy 4733 TCGTAGTTTATCTACACGACGGGGAGTCAGGCAACTATGATGAACGAATA 4783
```



```

|||||
923 TCGTAGTATCTACACGAGGGGAGTCAGGCAACTATGGATGACGAATAGA 973

RESULT 5
CD458286                                917 bp    mRNA    linear    EST 03-JUN-2003
LOCUS                                     Fg08_09a03_R Fg08_AAPC_ECORC_Fusarium_graminearum_complex_substrate
DEFINITION                               Gibberella zeae cDNA clone Fg08_09a03, mRNA sequence.
ACCESSION                               CD458286
VERSION
KEYWORDS
SOURCE
ORGANISM
Gibberella zeae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreales; Nectriaceae; Gibberella.
REFERENCE
1 (bases 1 to 917)
AUTHORS
Watson, R.J., Heyes, R., Chapados, J., Couroux, P., Harris, L.J.,
Hattori, J., Lacroix, C., Ouellet, T., Robert, L.S., Singh, J.A.,
Sprott, D. and Tinker, N.A.
TITLE
A cDNA library prepared from Fusarium graminearum grown on a
complex plant substrate
JOURNAL
Unpublished (2003)
COMMENT
Contact: Watson, Robert J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
Bldg. 20, Central Experimental Farm, Ottawa, Ontario, KIA 0C6,
CANADA
Tel: (613) 759-1655
Fax: (613) 759-1701
Email: watsonrj@agr.gc.ca.

FEATURES
source
1..917
/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="DAOM 180378"
/db_xref="taxon:5518"
/clone="Fg08_09a03"
/tissue_type="Mycelium"
/dev_stage="Asexual"
/lab_host="E. coli DH10B"
/clone_lib="Fg08_AAPC_ECORC_Fusarium_graminearum_complex_s
ubstrate"
/notes="vector: pBluescript II+; Site_1: EcoRI; Site_2:
XhoI; Fusarium graminearum grown on a complex plant
substrate-- wheat leaves treated to remove most of the low
molecular weight, water-soluble components."

ORIGIN
Query Match      15.2%; Score 885.6; DB 14; Length 917;
Best Local Similarity 99.6%; Pred. No. 3.2e-213;
Matches 886; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3833 ATACAGGTGGCACTTTTCGGGGAATGTGGCGGAACCCCTATTGTTTATTTTCTAAA 3892
DB 26 ATTTAGTGGGCACATTTTCGGGGAATGTGGCGGAACCCCTATTGTTTATTTTCTAAA 85
QY 3893 TACATTCAAATATGATCGCTCATGAGACATAACCCCTGATAAATGCTTCAATAATATT 3952
DB 86 TACATTCAAATATGATCGCTCATGAGACATAACCCCTGATAAATGCTTCAATAATATT 145
QY 3953 GAAAAAGAGAGATGATGATTCACATTTCCGTGTGCGCCCTATTCCCTTTTTCGGG 4012
DB 146 GAAAAAGAGAGATGATGATTCACATTTCCGTGTGCGCCCTATTCCCTTTTTCGGG 205
QY 4013 CATTTTCCTTCCTGTTTTCCTCACCAGAAAGCGTGGTGAAGTAAAGATGCTGAAG 4072
DB 206 CATTTTCCTTCCTGTTTTCCTCACCAGAAAGCGTGGTGAAGTAAAGATGCTGAAG 265
QY 4073 ATCAGTTGGGTGCACGAGTGGGTTTACATCGAAGTGGATCTCAACAGCGGTGAAGTCTTG 4132
DB 266 ATCAGTTGGGTGCACGAGTGGGTTTACATCGAAGTGGATCTCAACAGCGGTGAAGTCTTG 325

```

---

```

QY 4133 AGAGTTTTCGCCCGAAGACGTTTTTCAATGATGAGCACATTTTAAAGTTCTGCTATGTG 4192
DB 326 AGAGTTTTCGCCCGAAGACGTTTTTCAATGATGAGCACATTTTAAAGTTCTGCTATGTG 385
QY 4193 GCGCGGTATTATCCGTTATTGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATT 4252
DB 386 GCGCGGTATTATCCGTTATTGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATT 445
QY 4253 CTCAGAAATGACTTGGTTGAGTACTCACAGTGCACAGAAAAGCATCTTACGGATGGCATGA 4312
DB 446 CTCAGAAATGACTTGGTTGAGTACTCACAGTGCACAGAAAAGCATCTTACGGATGGCATGA 505
QY 4313 CAGTAAGAGAAATTATGCAAGTCTGCCATAACATGAGTGATAAACAATCGGGCCCACTTAC 4372
DB 506 CAGTAAGAGAAATTATGCAAGTCTGCCATAACATGAGTGATAAACAATCGGGCCCACTTAC 565
QY 4373 TTCTGACAAACGATCGGAGGACCGAAGAGCTAAACCGCTTTTTCGCAACAATCGGGGATC 4432
DB 566 TTCTGACAAACGATCGGAGGACCGAAGAGCTAAACCGCTTTTTCGCAACAATCGGGGATC 625
QY 4433 ATGTAACCTGCGCTTGTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGC 4492
DB 626 ATGTAACCTGCGCTTGTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGC 685
QY 4493 GTGACACACGATCGCTGTAGCAATGSCAACCAAGCTTGGCGCAAACTATTAACTGGCGAC 4552
DB 686 GTGACACACGATCGCTGTAGCAATGSCAACCAAGCTTGGCGCAAACTATTAACTGGCGAC 745
QY 4553 TACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGAGCGGAGGATGAAGTTGCAG 4612
DB 746 TACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGAGCGGAGGATGAAGTTGCAG 805
QY 4613 GACCATTCTGCGCTCGCGCCCTTCGGCTGCGCTGTTTATTTGCTGATAAATCTGAGCGG 4672
DB 806 GACCATTCTGCGCTCGCGCCCTTCGGCTGCGCTGTTTATTTGCTGATAAATCTGAGCGG 865
QY 4673 GTGAGCGTGGTCTCGCGGTATCATTCGACACTGGGCGCAGATGTTAAGCC 4724
DB 866 GTGAGCGTGGTCTCGCGGTATCATTCGACACTGGGCGCAGATGTTAAGCC 917

RESULT 6
BG838279                                935 bp    mRNA    linear    EST 25-MAY-2001
LOCUS                                     Gc01_10e07_R Gc01_AAPC_ECORC_cold stressed Glycine clandestina
DEFINITION                               Glycine clandestina cDNA clone Gc01_10e07, mRNA sequence.
ACCESSION                               BG838279
VERSION
KEYWORDS
SOURCE
Glycine clandestina
Glycine clandestina
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE
1 (bases 1 to 935)
AUTHORS
Singh, J.A., Farah, S., Chapados, J., Couroux, P., De Moors, A.,
Harris, L.J., Hattori, J.I., Ouellet, T., Robert, L.S., Sprott, D. and
Tinker, N.A.
TITLE
Expressed Sequence Tags from Cold-Stressed Glycine clandestina
Seedlings
JOURNAL
Unpublished (2001)
COMMENT
Contact: Singh, J.A.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA
0C6, Canada
Tel: (613) 759-1662
Fax: (613) 759-1701
Email: singhja@agr.gc.ca.
Location/Qualifiers
1..935
/organism="Glycine clandestina"

```

```
/mol_type="mRNA"
/cultivar="1035"
/db_xref="taxon:45687"
/clone="Gc01_10e07"
/tissue_type="Leaves, stem"
/clone_lib="Gc01_AAPC_ECORC_cold_stressed_glycine_clandest
ina"
/notes="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI;
Site 2: XhoI; Plants incubated at 2 degrees under 12 hours
of light/day. Harvested after only 2-3 days of cold
treatment. cDNA was prepared with the Uni-Zap cDNA kit
from Stratagene. Eco RI adapters were linked followed by
digest with Xho I/Eco RI and ligated to pBluescript."

ORIGIN
Query Match 15.0%; Score 873; DB 12; Length 935;
Best Local Similarity 97.9%; Pred. No. 5.1e-210;
Matches 878; Conservative 14; Mismatches 4; Indels 1; Gaps 1;

QY 3833 ATACAGTGGCGACTTTTCGGGGAATGTGGCGGAACCCCTATTTCGTTATTTCTTAA 3892
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
32 ATTTAGTGGCGACTTTTCGGGGAATGTGGCGGAACCCCTATTTCGTTATTTCTTAA 91

QY 3893 TACATTTCAATATGATTCGCTCATGAGACAATAACCTGATAAATGCTTCAATATATT 3952
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
92 TACATTTCAATATGATTCGCTCATGAGACAATAACCTGATAAATGCTTCAATATATT 151

QY 3953 GAAAAAGGAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTGGCG 4012
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
152 GAAAAAGGAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTGGCG 211

QY 4013 CATTTTGCCTTCTGTTTTTGTCTACCCGGAACCGCTGGTGAAGTAAAGATGCTGAAG 4072
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
212 CATTTTGCCTTCTGTTTTTGTCTACCCGGAACCGCTGGTGAAGTAAAGATGCTGAAG 271

QY 4073 ATCAGTTGGGTGCACAGTGGGTTCATCGAATCGATCTCAACAGCGTGAAGATCCTTG 4132
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
272 ATCAGTTGGGTGCACAGTGGGTTCATCGAATCGATCTCAACAGCGTGAAGATCCTTG 331

QY 4133 AGAGTTTTCGCCCCGGAAGACGTTTTTCCAATGATGAGCACTTTTAAAGTTTGTCTATGTG 4192
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332 AGAGTTTTCGCCCCGGAAGACGTTTTTCCAATGATGAGCACTTTTAAAGTTTGTCTATGTG 391

QY 4193 GCGCGTATATCCGTTATTCAGCGCGGCAAGAGCAATCGGTGCGCCGATACACTATT 4252
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
392 GCGCGTATATCCGTTATTCAGCGCGGCAAGAGCAATCGGTGCGCCGATACACTATT 451

QY 4253 CTCAGAAATGACTTGGTTGAGTACTCTACCAAGTACAGAAAAGCAATCTTACGGATGCGATGA 4312
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
452 CTCAGAAATGACTTGGTTGAGTACTCTACCAAGTACAGAAAAGCAATCTTACGGATGCGATGA 511

QY 4313 CAGTAAGAGAAATPATCAGTGTGCTCCATTAACCATGATGATACACTGCGG-CCAACTTA 4371
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
512 CAGTAAGAGAAATPATCAGTGTGCTCCATTAACCATGATGATACACTGCGGCGCCAACTTA 571

QY 4372 CTTCTGACAAACGATCCGAGGACCGAAGGAGCTAACCGCTTTTTCGACAAACATGGGGAT 4431
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
572 CTTCTGACAAACGATCCGAGGACCGAAGGAGCTAACCGCTTTTTCGACAAACATGGGGAT 631

QY 4432 CATGTAATCGCCTTGATCGTTGGGAACCGGAGCTGAATGAGCCATACCAACGACGAG 4491
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
632 CATGTAATCGCCTTGATCGTTGGGAACCGGAGCTGAATGAGCCATACCAACGACGAG 691

QY 4492 CGTGACACCAACGATGCTGTAGCAATGGCAACAAACGTTGGCGCAACTTATTAACCTGGCGAA 4551
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
692 CGTGACACCAACGATGCTGTAGCAATGGCAACAAACGTTGGCGCAACTTATTAACCTGGCGRA 751

QY 4552 CTACTTACTCTAGCTTCCCGGCAACAAATTAATAGACTGAGTGGCGGCGATTAAGTTGCA 4611
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
752 CTACTTACTCTAGCTTCCCGGCAACAAATTAATAGACTGAGTGGCGGCGATTAAGTTGCA 811

QY 4612 GGACCACTTCTGGCTCGGCCCTTCGCGCTGGCTGTTATTGCTGATAAATCTCGAGGC 4671
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
DB 812 GGACCACTTCTGGCTCGGCCCTTCGCGTGGTGGTATTATTGCTGATAWATCTGGGCC 871
QY 4672 GGTGACGTGGGTCTCGCGGTATCATTTGAGCACTGGGCGCAGATGTAAGCCCTCC 4728
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
872 GGTGACGTGGGTCTCGCGGTATCATTTGAGCACTGGGCGCAGATGTAAGCCCTCC 928

RESULT 7
CD459092 918 bp mRNA linear EST 03-JUN-2003
LOCUS Fg08_08e02_R Fg08 AAPC_ECORC_Fusarium_graminearum_complex_substrate
DEFINITION Gibberella zeae cDNA clone Fg08_08e02, mRNA sequence.
ACCESSION CD459092
VERSION CD459092.1 GI:31373832
KEYWORDS EST.
SOURCE Gibberella zeae
ORGANISM Gibberella zeae
Bukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE 1 (bases 1 to 918)
AUTHORS Watson,R.J., Heys,R., Chapados,J., Couroux,P., Harris,L.J.,
Hattori,J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A.,
Sprott,D. and Tinker,N.A.
TITLE A cDNA library prepared from Fusarium graminearum grown on a
JOURNAL complex plant substrate
COMMENT Unpublished (2003)
Contact: Watson, Robert.J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
Bldg. 20, Central Experimental Farm, Ottawa, Ontario, KIA 0C6,
CANADA
Tel: (613) 759-1655
Fax: (613) 759-1701
Email: watsonrj@agr.gc.ca.

FEATURES
Location/Qualifiers
source
1..918
/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="DAOM 180378"
/db_xref="taxon:5518"
/clone="Fg08_08e02"
/tissue_type="Mycelium"
/dev_stage="Asexual"
/lab_host="E. coli DH10B"
/clone_lib="Fg08_AAPC_ECORC_Fusarium_graminearum_complex_s
ubstrate"
/notes="Vector: pBluescript II+; Site_1: EcoRI; Site_2:
XhoI; Fusarium graminearum grown on a complex plant
substrate-- wheat leaves treated to remove most of the low
molecular weight, water-soluble components."

ORIGIN
Query Match 14.9%; Score 865.4; DB 14; Length 918;
Best Local Similarity 99.2%; Pred. No. 4.4e-208;
Matches 888; Conservative 2; Mismatches 3; Indels 2; Gaps 2;

QY 3833 ATACAGTGGCGACTTTTCGGGGAATGTGGCGGAACCCCTATTTCGTTATTTCTTAA 3892
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 ATTTAGTGGCGACTTTTCGGGGAATGTGGCGGAACCCCTATTTCGTTATTTCTTAA 83

QY 3893 TACATTTCAATATGATTCGCTCATGAGACAATAACCTGATAAATGCTTCAATATATT 3952
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
84 TACATTTCAATATGATTCGCTCATGAGACAATAACCTGATAAATGCTTCAATATATT 143

QY 3953 GAAAAAGGAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTGGCG 4012
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
144 GAAAAAGGAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTGGCG 203

QY 4013 CATTTTGCCTTCTGTTTTTGTCTACCCGGAACCGCTGGTGAAGTAAAGATGCTGAAG 4072
DB |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
204 CATTTTGCCTTCTGTTTTTGTCTACCCGGAACCGCTGGTGAAGTAAAGATGCTGAAG 263

QY 4073 ATCAGTTGGGTGCACGAGTGGGTTCATCGAATCGATCTCAACAGCGGTGAAGATCCTTG 4132
```

Db	264	ATCAGTTGGGTGACAGTGGTGGTTACATCGAACCTGGATCTCAACAGCGGTAAAGATCTTGT	323
Qy	4133	AGAGTGTTCGCCCGGAGAAAGTTCCTCAATGATGAGCAGCTTTTAAAGTTCCTGATGTG	4192
Db	324	AGAGTGTTCGCCCGGAGAAAGTTCCTCAATGATGAGCAGCTTTTAAAGTTCCTGATGTG	383
Qy	4193	GCGCGGTATTATCCCGTATTGACCGCGGGCAAGAGCACTCGGTGCGCGCATACACTATT	4252
Db	384	GCGCGGTATTATCCCGTATTGACCGCGGGCAAGAGCACTCGGTGCGCGCATACACTATT	443
Qy	4253	CTCAGATGACTTGGTGTGAGTACTCACCAGTTCACAGAAAGCATCTTACGGATGGCATGA	4312
Db	444	CTCAGATGACTTGGTGTGAGTACTCACCAGTTCACAGAAAGCATCTTACGGATGGCATGA	503
Qy	4313	CAGTAAGAGAAATTATGCAAGTCTGCCATAACCA-TGAGTGATACACTCGCGGCCAATTCTA	4371
Db	504	CAGTAAGAGAAATTATGCAAGTCTGCCATAACCA-TGAGTGATACACTCGCGGCCAATTCTA	563
Qy	4372	CTTCTGACACAGATCGGAGGACCGAAGAGCTAACCGGTTTTTGGACACATGGGGAT	4431
Db	564	CTTCTGACACAGATCGGAGGACCGAAGAGCTAACCGGTTTTTGGACACATGGGGAT	623
Qy	4432	CATGTAACCTGCGCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAGCAGAG	4491
Db	624	CATGTAACCTGCGCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAGCAGAG	683
Qy	4492	CGTGACACACAGATCGCTGTAGCAATGGCAACAACTGTTGGCAAACTATTAACTGGCGAA	4551
Db	684	CGTGACACACAGATCGCTGTAGCAATGGCAACAACTGTTGGCAAACTATTAACTGGCGAA	743
Qy	4552	CTACTTACTTACTGCTTCGGGCAACAAATTAATAGACTGGAGGCGGATTAAGTTGCA	4611
Db	744	CTACTTACTTACTGCTTCGGGCAACAAATTAATAGACTGGAGGCGGATTAAGTTGCA	803
Qy	4612	GGACCACTCTGGCTCGGCCCTTCGCGCTGGCTGTTTATGCTGATTAATCTGGAGCC	4671
Db	804	GGACCACTCTGGCTCGGCCCTTCGCGCTGGCTGTTTATGCTGATTAATCTGGAGCC	863
Qy	4672	GGTGAGCGTGGGTCTCGCGGT-ATCATTGACGACACTGGGGCCAGATGGTAAGGCC	4725
Db	864	GGTGAGCGTGGGTCTCGCGGT-ATCATTGACGACACTGGGGCCAGATGGTAAGGCC	918
RESULT 8			
AL042026			
LOCUS		841 bp mRNA linear EST 04-SEP-2003	
DEFINITION		DKFZp434E111 r1 434 (synonym: htes3) Homo sapiens cDNA clone	
ACCESSION		DKFZp434E111.5, mRNA sequence.	
VERSION		AL042026.1 GI:5421372	
KEYWORDS		EST.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE		1 (bases 1 to 841)	
JOURNAL		Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.	
COMMENT		EST (Poustka, et al.)	
		Unpublished (1999)	
		Contact: MIPS	
		MIPS	
		Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany	
		This is the 5' sequence of the clone insert	
		Clone from S. Wiemann, Molecular Genome Analysis, German Cancer	
		Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;	
		sequenced by DKFZ (German Cancer Research Center,	
		Heidelberg/Germany) within the cDNA sequencing consortium of the	
		German Genome Project.	
		No 51 sequence available.	
		This clone (DKFZp434E111) is available at the RZPD in Berlin.	
		Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059	
		Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.	

FEATURES	Location/Qualifiers
source	1..841
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="DKFZp434E111"
	/tissue_type="testis"
	/dev_stage="adult"
	/lab_host="DH10B"
	/clone_lib="434 (synonym: htes3)"
	/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
ORIGIN	
Query Match	14.4%; Score 841; DB 9; Length 841;
Best Local Similarity	100.0%; Pred. No. 6.7e-202;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	3873 TATTGTGTTTATTTTCTAAATACATTCAAATATGATGATCGGTCTATGACACATAACCTGTG
Db	1 TATTGTGTTTATTTTCTAAATACATTCAAATATGATGATCGGTCTATGACACATAACCTGTG
Qy	3933 ATAAATGCTTCAATATATTTGAAAAGGAGAGATATGAGTATTCACAACTTTCCGTGTGCG
Db	61 ATAAATGCTTCAATATATTTGAAAAGGAGAGATATGAGTATTCACAACTTTCCGTGTGCG
Qy	3993 CTTTATTCCTTTTGGGGCATTTTGGCTTCCTGTTTCTGCTCACCAGAAACGCTGCT
Db	121 CTTTATTCCTTTTGGGGCATTTTGGCTTCCTGTTTCTGCTCACCAGAAACGCTGCT
Qy	4053 GAAAGTAAAGATGCTCAAGATCAGTTGGGTGACAGAGTGGTTTACATCGAACTGGAATCT
Db	181 GAAAGTAAAGATGCTCAAGATCAGTTGGGTGACAGAGTGGTTTACATCGAACTGGAATCT
Qy	4113 CAACAGCGGTAAAGATCTTCAGAGTCTTTCGCCCGGAGAGAAAGTTCATATGATGAGCAC
Db	241 CAACAGCGGTAAAGATCTTCAGAGTCTTTCGCCCGGAGAGAAAGTTCATATGATGAGCAC
Qy	4173 TTTTAAAGTTCTGCTATGTCGCGGTATTTATCCGTTATGACCGCGGGCAAGCAACT
Db	301 TTTTAAAGTTCTGCTATGTCGCGGTATTTATCCGTTATGACCGCGGGCAAGCAACT
Qy	4233 CGGTGCGCGCATACACTATTCTCAGAAATGACTTGGTTGAGTACTCACCAGTCAACAGAAA
Db	361 CGGTGCGCGCATACACTATTCTCAGAAATGACTTGGTTGAGTACTCACCAGTCAACAGAAA
Qy	4293 GCATCTTACCGATGGCATGACAGTAAGAGAAATATGCACTGCTGCCATAACCATGAGTGA
Db	421 GCATCTTACCGATGGCATGACAGTAAGAGAAATATGCACTGCTGCCATAACCATGAGTGA
Qy	4353 TAACACTGCGGCCCACTTACTCTCTGACAAAGATCGGAGGACCGAGGAGCTTAACCGCTTT
Db	481 TAACACTGCGGCCCACTTACTCTCTGACAAAGATCGGAGGACCGAGGAGCTTAACCGCTTT
Qy	4413 TTTTGCAACAACTGGGGGATCATGTAACCTCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGT
Db	541 TTTTGCAACAACTGGGGGATCATGTAACCTCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGT
Qy	4473 AGCCATACCAAAAGTTCGAGGACCACTTCTGCGCTCGCGCCCTTCGGCTGCTGCTTAT
Db	601 AGCCATACCAAAAGTTCGAGGACCACTTCTGCGCTCGCGCCCTTCGGCTGCTGCTTAT
Qy	4533 CAATCTTAACTGGCGGAACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT
Db	661 CAATCTTAACTGGCGGAACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT
Qy	4593 GAGGCGGATTAAGTTCGAGGACCACTTCTGCGCTCGCGCCCTTCGGCTGCTGCTTAT
Db	721 GAGGCGGATTAAGTTCGAGGACCACTTCTGCGCTCGCGCCCTTCGGCTGCTGCTTAT
Qy	4653 TGCTGATAAATCTGGAGCGGTGAGGTGCTGCGGGTATCATTCAGACACTGGGGCC
Db	781 TGCTGATAAATCTGGAGCGGTGAGGTGCTGCGGGTATCATTCAGACACTGGGGCC

```

QY      4713 A 4713
Db      841 A 841

RESULT 9
CD459085
LOCUS   Fg08_08d03_R_Fg08_AAFC_ECORC_Fusarium_graminearum_complex_substrate
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
Hattori,J., Heyes,R., Chapados,J., Couroux,P., Harris,L.J.,
Sprott,D. and Tinker,N.A.
TITLE
A cDNA library prepared from Fusarium graminearum grown on a
complex plant substrate
JOURNAL
COMMENT
Contact: Watson, Robert.J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
Bldg. 20, Central Experimental Farm, Ottawa, Ontario, KIA 0C6,
CANADA
Tel: (613) 759-1655
Fax: (613) 759-1701
Email: watsonrj@agr.gc.ca.

FEATURES
source
1..872
/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="DAOM 180378"
/db_xref="taxon:5518"
/clone="Fg08_08d03"
/tissue_type="Mycelium"
/dev_stage="Asexual"
/lab_host="E. coli DH10B"
/clone_lib="Fg08_AAFC_ECORC_Fusarium_graminearum_complex_s
ubstrate"
/note="vector: pluescript II+; Site_1: EcoRI; Site_2:
XhoI; Fusarium graminearum grown on a complex plant
substrate-- wheat leaves treated to remove most of the low
molecular weight, water-soluble components."

ORIGIN
Query Match 14.38; Score 832.8; DB 14; Length 872;
Best Local Similarity 99.24; Pred. No. 8.1e-200;
Matches 831; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      3846 TTTTCGGGGAAATGTGCGCGGAACCCCTATTGTTTATTTTCTAAATACATTCACAAATAT
Db      35 TTTTCGGGGAAATGTGCGCGGAACCCCTATTGTTTATTTCTAAATACATTCACAAATAT 94

QY      3906 GTATCCGCTCATGAGACAAATACCTGATTAATGCTTCAATATATTGAAAGGAAGAG
Db      95 GTATCCGCTCATGAGACAAATACCTGATTAATGCTTCAATATATTGAAAGGAAGAG 154

QY      3966 TATGAGTATTCAACATTTCCGTGCGCCCTATTCCCTTTTTCGGGCATTTTCCTTCC
Db      155 TATGAGTATTCAACATTTCCGTGCGCCCTATTCCCTTTTTCGGGCATTTTCCTTCC 214

QY      4026 TGTTTTGTCTCACCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGC
Db      215 TGTTTTGTCTCACCAGAGAGCGTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGC 274

QY      4086 ACGATGGGTGTACATCAACTGATCTCAACAGCGGTGAAGATCTTCAGAGATTTTCGCC
Db      275 ACGATGGGTGTACATCAACTGATCTCAACAGCGGTGAAGATCTTCAGAGATTTTCGCC 334

```

```

QY      4146 CGAAGAACGTTTTTCCAATGATGAGCACATTTTAAAGTTCTGCTATGTGGCGGCGGTATTATC
Db      335 CGAAGAACGTTTTTCCAATGATGATGAGCACATTTTAAAGTTCTGCTATGTGGCGGCGGTATTATC 394

QY      4206 CCGTATTGACCGCGCGGCAAGAGCAACTCGGTGCGGCATACACTATTCTCAGAATGACTT
Db      395 CCGTATTGACCGCGCGGCAAGAGCAACTCGGTGCGGCATACACTATTCTCAGAATGACTT 454

QY      4266 GCTTGAGTACTTCACCACTGACAGAAAGCATCTTACCGATGGCATGACAGTAAGAGAAAT
Db      455 GCTTGAGTACTTCACCACTGACAGAAAGCATCTTACCGATGGCATGACAGTAAGAGAAAT 514

QY      4326 ATGCAGTGTGCCATACCACTAGTGAATACACTGCGGCGCAACTTACTTCTGACAAACGAT
Db      515 ATGCAGTGTGCCATACCACTAGTGAATACACTGCGGCGCAACTTACTTCTGACAAACGAT 574

QY      4386 CGGAGGACCGAAGGAGCTAAACCGCTTTTTCGCAACAATGGGGGATCANTGTAATCTGCGCT
Db      575 CGGAGGACCGAAGGAGCTAAACCGCTTTTTCGCAACAATGGGGGATCANTGTAATCTGCGCT 634

QY      4446 TGATCGTTGGACCGCGAGCTGAATGAAGCCATACCAACGAGCGGTGACACCAGAT
Db      635 TGATCGTTGGACCGCGAGCTGAATGAAGCCATACCAACGAGCGGTGACACCAGAT 694

QY      4506 GCCTGTAGCAATGCAACAACGTTGCGCAAACTATTAACTGCGCAACTTACTTCTCTAGC
Db      695 GCCTGTAGCAATGCAACAACGTTGCGCAAACTATTAACTTAACTGCGCAACTTACTTCTCTAGC 754

QY      4566 TTCCCGGCAACAATTAATAGACTGGATGGAGCGCGGATAAAGTTTCAGAGCACTTCTCGG
Db      755 TTCCCGGCAACAATTAATAGACTGGATGGAGCGCGGATAAAGTTTCAGAGCACTTCTCGG 814

QY      4626 CTGCGGCCCTTCGCGCTGGCTGTTATTGCTGATAAATCTGGAGCCGCTGAGCGGTGGG
Db      815 CTGCGGCCCTTCGCGCTGGCTGTTATTGCTGATAAATCTGGAGCCGCTGAGCGGTGGG 872

RESULT 10
LOCUS   CP269652/c
DEFINITION
CP269652 1073 bp mRNA linear EST 13-AUG-2003
Fgylcolid844 Fragilariopsis cylindrus SMART cDNA library (Clontech)
Fragilariopsis cylindrus cDNA clone Antarctic 5', mRNA sequence.
CP269652
VERSION
CP269652.1 GI:33631539
KEYWORDS
EST.
SOURCE
Fragilariopsis cylindrus
Fragilariopsis cylindrus
Bakaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
Bacillariophycidae; Bacillariales; Bacillariaceae; Fragilariopsis.
1 (bases 1 to 1073)
Mock,T. and Valentin,K.
REFERENCE
1. Mock,T. and Valentin,K.
AUTHORS
EST analysis of freezing tolerance in the Antarctic diatom
Fragilariopsis cylindrus: Detection of numerous cold adaptation
related genes and gene transfer events
JOURNAL
Unpublished (2003)
COMMENT
Contact: Mock T
Biological Oceanography
Alfred-Wegener-Institute for Polar and Marine Research
Am Handelshafen 12, D-27570 Bremerhaven, Germany
Tel: +49 471 4831 1893
Fax: +49 471 4831 1425
Email: tmock@awi-bremerhaven.de
sequence with unknown function
PCR Primers
FORWARD: 5'lambdaTriplex2
BACKWARD: 3'lambdaTriplex2
Seq primer: ctcgggaagcgcgcattgtgtggt.
FEATURES
source
1..1073
/organism="Fragilariopsis cylindrus"
/mol_type="mRNA"
/strain="Antarctic"

```



```
Db 301 GACAGTAAGAGAAATATGCGATGCTGCATTAACCATGAGTGATTAACATCTGGCCCAACTT 360
Qy 4371 ACTTCTGACCAACGATCGGAGGACCGAAGAGCTTAACCGCTTTTGTGCAACATCTGGGGA 4430
Db 361 ACTTCTGACCAACGATCGGAGGACCGAA-GAGCTAAACCGCTTTTGTGCAACATCTGGGGA 419
Qy 4431 TCATGTAACCTGCTTGTATGCTTGGGAAACCGGAGCTGAATGAAGCCATACCAACGACGA 4490
Db 420 TCATGTAACCTGCTTGTATGCTTGGGAAACCGGAGCTGAATGAAGCCATACCAACGACGA 479
Qy 4491 GCGTGACACCAACGATGCTGTAGCAATGGCAACGTTGGCAAACTATTAACTGCGGA 4550
Db 480 GCGTGACACCAACGATGCTGTAGCAATGGCAACGTTGGCAAACTATTAACTGCGGA 539
Qy 4551 ACTACTTACTCTAGCTTCCCGGCAACAAATTAATAGACTGGATGGAGCGGATAAGTTGC 4610
Db 540 ACTACTTACTCTAGCTTCCCGGCAACAAATTAATAGACTGGATGGAGCGGATAAGTTGC 599
Qy 4611 AGGACCACTTCTGGCTCGGCCCTTCCGGCTGGCTGGCTGGTTTATTGCTGATAAATCTGGAGC 4670
Db 600 AGGACCACTTCTGGCTCGGCCCTTCCGGCTGGCTGGCTGGTTTATTGCTGATAAATCTGGAGC 659
Qy 4671 CGGTGAGCGTGGTCTCGCGGTATCATTTGACGACCTGGGCGCAGATGTAAGCCCTCCCG 4730
Db 660 CGGTGAGCGTGGTCTCGCGGTATCATTTGACGACCTGGGCGCAGATGTAAGCCCTCCCG 719
Qy 4731 TATCTAGTATTATCTACACGACGGGAGTCAAGCACTATGATGATGAACGAAATAGACAGAT 4790
Db 720 TATCTAGTATTATCTACACGACGGGAGTCAAGCACTATGATGATGAACGAGATAGACAGAT 779
Qy 4791 CGGTGATAGTGGCTCTACATGATTAAGCAATGATTAAGCAATGATTAAGCAATGATTAAGCA 4850
Db 780 CGGTGATAGTGGCTCTACATGATTAAGCAATGATTAAGCAATGATTAAGCAATGATTAAGCA 839
Qy 4851 TATACTTTAGATTGA 4865
Db 840 TATGTTGAGATTGA 854
```

```
RESULT 12
LOCUS CD458333
DEFINITION Fg08_09e07_R Fg08_AAPC_ECORC_Fusarium_graminearum_complex_substrate
Gibberella zeae cDNA clone Fg08_09e07, mRNA sequence.
ACCESSION CD458333
VERSION CD458333.1 GI:31373073
KEYWORDS EST.
SOURCE Gibberella zeae
ORGANISM Gibberella zeae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
1 (bases 1 to 870)
Watson, R.J., Heyes, R., Chapados, J., Couroux, P., Harris, L.J.,
Hattori, J., Lacroix, C., Ouellet, T., Robert, L.S., Singh, J.A.,
Sprott, D., and Tinker, N.A.
A cDNA library prepared from Fusarium graminearum grown on a
complex plant substrate
Unpublished (2003)
Contact: Watson, Robert J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-Food Canada
Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA
Tel: (613) 759-1655
Fax: (613) 759-1701
Email: watsonrj@agr.gc.ca.
Location/Qualifiers
1. 870
/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="DAOM 180378"
/db_xref="taxon:5518"
```

## FEATURES

## source

```
FEATURES
source
1. 870
/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="DAOM 180378"
/db_xref="taxon:5518"
```

```
/clone="Fg08_09e07"
/tissue_type="Mycelium"
/dev_stage="Asexual"
/lab_host="E. coli DH10B"
/clone_lib="Fg08_AAPC_ECORC_Fusarium_graminearum_complex_s
ubstrate"
/notes="Vector: pBluescript II+; Site_1: EcoRI; Site_2:
XhoI; Fusarium graminearum grown on a complex plant
substrate-- wheat leaves treated to remove most of the low
molecular weight, water-soluble components."
```

## ORIGIN

```
Query Match 14.18; Score 819.6; DB 14; Length 870;
Best Local Similarity 99.2%; Pred. No. 1.8e-196;
Matches 841; Conservative 3; Mismatches 2; Indels 2; Gaps 2;
Qy 3833 ATACAGGTGGCAGCTTTTCGGGGAATGTCGGCGAAACCCCTATTGTTGTTATTTTCTAAA 3892
Db 24 ATTAGGTGGCASTTTTCGGGGAATGTCGGCGAAACCCCTATTGTTATTTTCTAAA 83
Qy 3893 TACATTCAAAATATGATCCCTCATGAGACAATTAACCCCTGATAAATGCTTCAATAATTT 3952
Db 84 TACATTCAAAATATGATCCCTCATGAGACAATTAACCCCTGATAAATGCTTCAATAATTT 143
Qy 3953 GAAAAGGAAGATGATGATTTCAACATTTCCGTGTGCGCCCTATTCCCTTTTTCGGG 4012
Db 144 GAAAAGGAAGATGATGATTTCAACATTTCCGTGTGCGCCCTATTCCCTTTTTCGGG 203
Qy 4013 CATTTTGCCTTCCCTTTTTCCTCACCACAGAAACGCTGTGAAAGTAAAGATGCTGAAG 4072
Db 204 CATTTTGCCTTCCCTTTTTCCTCACCACAGAAACGCTGTGAAAGTAAAGATGCTGAAG 263
Qy 4073 ATCAGTTGGTGAACGAGTCGGTTTACATCGAATCGGATCTCAACAGCGGTGAAGATCTT 4132
Db 264 ATCAGTTGGTGAACGAGTCGGTTTACATCGAATCGGATCTCAACAGCGGTGAAGATCTT 323
Qy 4133 AGAGTTTTCGCCCCGGAAGAACGTTTTCATATGATGACACATTTTAAAGTTCGTATGTG 4192
Db 324 AGAGTTTTCGCCCCGGAAGAACGTTTTCATATGATGACACATTTTAAAGTTCGTATGTG 383
Qy 4193 GCGCGGTATTATCCCGTATTGAGCGCGGGAAGAGCAATCGGTGCGCGCATACATATT 4252
Db 384 GCGCGGTATTATCCCGTATTGAGCGCGGGAAGAGCAATCGGTGCGCGCATACATATT 443
Qy 4253 CTGAGTAATGATCTGGTTGAGTACTCAACAGTCAAGAAAAGCATCTTACGATGCGCATGA 4312
Db 444 CTGAGTAATGATCTGGTTGAGTACTCAACAGTCAAGAAAAGCATCTTACGATGCGCATGA 503
Qy 4313 CAGTAAGGAATTAATGACAGTGTGCCATTAACCATGAGTGAATAACACTGCGGCCAATTTAC 4372
Db 504 CAGTAAGGAATTAATGACAGTGTGCCATTAACCATGAGTGAATAACACTGCGGCCAATTTAC 563
Qy 4373 -TTCTGACAAACGATCGGAGGACCGAAGAGCTAACCGCTTTTTCGCAACAATCGGGGAT 4431
Db 564 TTTCTGCAACGATCGGAGGACCGAA-GAGCTAAACCGCTTTTTCGCAACAATCGGGGAT 622
Qy 4432 CATGTAATCTGCTTGTATGCTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAG 4491
Db 623 CATGTAATCTGCTTGTATGCTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAG 682
Qy 4492 CGTGACACCAACGATGCTGTAGCAATGGCAACACGTTGGCAAACTATTAACTTGGCGAA 4551
Db 683 CGTGACACCAACGATGCTGTAGCAATGGCAACACGTTGGCAAACTATTAACTTGGCGAA 742
Qy 4552 CTACTTTACTCTAGCTTCCCGGCAACAAATTAATAGACTGGATGGAGCGGATAAGTTGCA 4611
Db 743 CTACTTTACTCTAGCTTCCCGGCAACAAATTAATAGACTGGATGGAGCGGATAAGTTGCA 802
Qy 4612 GGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATTAATCTGGAGCC 4671
Db 803 GGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATTAATCTGGAGCC 862
Qy 4672 GGTGAGCG 4679
```

```
Db 863 GKGAGCG 870
|||||
RESULT 13
BZ572284/c 1249 bp DNA linear GSS 17-DEC-2002
LOCUS msh2_2572.x3 msh Pseudomonas aeruginosa genomic clone msh2_2572,
DEFINITION genomic survey sequence.
ACCESSION BZ572284
VERSION BZ572284.1 GI:27207345
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1249)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers
source 1..1249
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clones="msh2_2572"
/clone_lib="msh"
/note="Environmental isolate. Whole genomic shotgun
library."

ORIGIN
Query Match 14.0%; Score 815.6; DB 28; Length 1249;
Best Local Similarity 93.0%; Pred. No. 2e-195;
Matches 952; Conservative 0; Mismatches 61; Indels 11; Gaps 9;

Qy 4799 TAGTGCTCACTGATTAAAGCAATGGTAAGTGTGAGCAAGTTTACTCATATATACTTT 4858
Db 1116 TATGGGCCCCCATATATAAACACTGGGACCGGACGACCAAGTTTCCCA-AAATACCTT 1058
Qy 4859 AGATTGATTAAAACTTCATTTTAAATTTAAAGGATCTAGTGAAGATCCCTTTTGATA 4918
Db 1057 AAATTGATTAAAACTTCC-TTTTAAATTTAAAGGATTAAAGT-AAAGATCCTTTTAAAA 1000
Qy 4919 ATCTCATGACCAAAATCCCTTAAAGTGAGTTTCTGTCACCTGAGCGTCAGACCCGCTAG 4978
Db 999 ATCTCATG--CCAAATCCCTTACGTGAGTTTCTGTCCTCACTAAGGGCCAGACCCGCTAA 942
Qy 4979 AAAAGATCAAAAGGATCTTCTTGAGATCCTTTTCTGCGGTAAATCTGCTGTGCAAA 5038
Db 941 GAAAGACAAAGATCTTCTGAGAACCTTTTCTGCGGTAAAT-TGCTGCTTGC-AA 884
Qy 5039 CAAAAAACCAACCGCTACAGCGGTGTTTGTTCGCGGATCAAGAGTACCAATCTTTT 5098
Db 883 CAAAAAACCAACCGCTACCAACCGTGGTTG--TTTGGCGATCAAGAGCTA-CAATCTTTT 827
Qy 5099 TTCCGAGGTAACCTGGCTTACGACGAGCGGACATACAAATCTGCTCTAGTGATGC 5158
Db 826 TTCGAGAGGTANCT-GCTTTCAGCAGAGCGGACATACAAATCTGTTCTTAGTGATGC 768
Qy 5159 CGTAGTAGGCCACCACTTCAAGAACTCTGTAGCAGCCGCTACATACCTCGCTCTGCTAA 5218
Db 767 CGTAGTAGGCCACCACTTCAAGAACTCTGTAGCAGCCGCTACATACCTCGCTCTGCTAA 708

FEATURES
source
Location/Qualifiers
```



```

source
1. .1126
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/db_xref="MSH"
/clone="mah2_5533"
/clone_lib="mah"
/note="Environmental isolate. Whole genomic shotgun
library."

ORIGIN
Query Match 14.08; Score 814.4; DB 28; Length 1126;
Best Local Similarity 90.4%; Pred. No. 4e-195;
Matches 938; Conservative 0; Mismatches 91; Indels 9; Gaps 6;

Qy 3881 TATTTTTCTAAATACATTCAAATATGATCGCTCATGAGACAAATACCCCTGATAAATGC 3940
Db 93 TATTTTTCTAAATACATTCAAATATGATCGCTCATGAGACAAATACCCCTGATAAATGC 152

Qy 3941 TTCAATAATATTTGAAAAGGAGAGTATGAGTATTTCAACATTTTCGGTGCGCCCTTATTC 4000
Db 153 TTCAATAATATTTGAAAAGGAGAGTATGAGTATTTCAACATTTTCGGTGCGCCCTTATTC 212

Qy 4001 CCTTTTTTGGCGCATTTTGGCTTCTCTGTTTTTGTCTCAACCAGAAAACGCTGGTGAAGTAA 4060
Db 213 CCTTTTTTGGCGCATTTTGGCTTCTCTGTTTTTGTCTCAACCAGAAAACGCTGGTGAAGTAA 272

Qy 4061 AAGATGCTGAAGATCAGTTCGGTGCAAGAGTGGGTATCATTCGAATCTCGATCTCAACAGCG 4120
Db 273 AAGATGCTGAAGATCAGTTCGGTGCAAGAGTGGGTATCATTCGAATCTCGATCTCAACAGCG 332

Qy 4121 GTAAGATCCTTTGAGAGTTTTTCGCCCCCGAAGAACGTTTTTCCAATGATCAGACATTTTAAAG 4180
Db 333 GTAAGATCCTTTGAGAGTTTTTCGCCCCCGAAGAACGTTTTTCCAATGATCAGACATTTTAAAG 392

Qy 4181 TTCGTCTATGTGGCGCGGTATTAATCCCGTATTTGACGCGCGGCAAGACAACTCGGTGCGC 4240
Db 393 TTCGTCTATGTGGCGCGGTATTAATCCCGTATTTGACGCGCGGCAAGACAACTCGGTGCGC 452

Qy 4241 GCATACACTATCTCAGAAAGTACTTGGTTCAGTACTCACCAAGTACACAGAAAGCATCTTA 4300
Db 453 GCATACACTATCTCAGAAAGTACTTGGTTCAGTACTCACCAAGTACACAGAAAGCATCTTA 512

Qy 4301 CGGATGGCATGACAGTAAGAAATTAATGCACTGCTGCCATAACCATGAGTGAATAACACTG 4360
Db 513 CGGATGGCATGACAGTAAGAAATTAATGCACTGCTGCCATAACCATGAGTGAATAACACTG 572

Qy 4361 CGGCCAACTTACTTCTGACAAAGATCGGAGGACCGAAGGAGCTAAACGCTTTTTTGTGCACA 4420
Db 573 CGGTCAACTTACTTATGACAAAGATCGGAGGACCGAAGGAGCTAAACGCTTTTTTGTGCACA 632

Qy 4421 ACATGGGGATCATGTAACCTCGCTTGATCGTTGGGAACCGGAGCTGAATGAAGACCATAC 4480
Db 633 ACATGGCGGATCATGTAACCTCGCTTGATCGTTGGGAACCGGAGCTGAATGAAGACCATAC 692

Qy 4481 CAAACGACGAGCGTGACACCAACGATGCTGTAGCAATGGCAACAAACGTCGCGCAAACTAT 4540
Db 693 CAAACGACGAGCGTGACACCAACGATGCTGTAGCAATGGCAACAAACGTCGCGCAACTAT 752

Qy 4541 TAACTGCGCAACTACTTAATCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGCGG 4600
Db 753 TAACTGCGCAACTACTGACTAGCTTTCGCGGACACAAATAATAGACTGGATGGAGCGGCTG 812

Qy 4601 AT-AAAGTTTCAGGACCACTTCTGGCTCGGCCCTTCGGCTGGCTGGTGTATTGCTGAT 4659
Db 813 ATAAAGTTTCAGGACCACTTTTGGGCTCGGCCCTTTCGGATGGCTGGTGTATTGTTGGA 872

Qy 4660 AAATCTCGAGCGCGTGAGCGTGGGTCTCGCGGTATCAATTCGACACTGGGGCGCAGATGGT 4719
Db 873 TAAATTCGAGCGCGGGGAAACGGGGGGCTTCGGGTTACATTGCACTACTTGGGGCAATAGG 932

Qy 4720 AAGCCCTCCCGTATCGTAGTTATCTACACGA-CGGGGAGTCAGGCAACTTA-TGGATGAAC 4777

```

